HIGHWAY RESEARCH RECORD

Number 383

Transportation Costs

5 reports

HIGHWAY RESEARCH BOARD

NATIONAL RESEARCH COUNCIL

NATIONAL ACADEMY OF SCIENCES—NATIONAL ACADEMY OF ENGINEERING

HIGHWAY RESEARCH RECORD

Number 383 | Transportation Costs

5 reports
prepared for the
51st Annual Meeting

Subject Areas

15 Transportation Economics
84 Urban Transportation Systems

HIGHWAY RESEARCH BOARD

DIVISION OF ENGINEERING NATIONAL RESEARCH COUNCIL
NATIONAL ACADEMY OF SCIENCES—NATIONAL ACADEMY OF ENGINEERING

Washington, D.C. 1972

NOTICE

The studies reported herein were not undertaken under the aegis of the National Academy of Sciences or the National Research Council. The papers report research work of the authors done at the institution named by the authors. The papers were offered to the Highway Research Board of the National Research Council for publication and are published herein in the interest of the dissemination of information from research, one of the major functions of the HRB.

Before publication, each paper was reviewed by members of the HRB committee named as its sponsor and was accepted as objective, useful, and suitable for publication by NRC. The members of the committee were selected for their individual scholarly competence and judgment, with due consideration for the balance and breadth of disciplines. Responsibility for the publication of these reports rests with the sponsoring committee; however, the opinions and conclusions expressed in the reports are those of the individual authors and not necessarily those of the sponsoring committee, the HRB, or the NRC.

Although these reports are not submitted for approval to the Academy membership or to the Council of the Academy, each report is reviewed and processed according to procedures established and monitored by the Academy's Report Review Committee.

ISBN 0-309-02052-2

Price: $2.00

Available from

Highway Research Board
National Academy of Sciences
2101 Constitution Avenue, N.W.
Washington, D.C. 20418

CONTENTS

FOREWORD

The technique of economic analysis most heavily relied on for transportation investment has been cost-benefit analysis. Over the years this procedure has become quite refined and sophisticated. In this RECORD, three papers deal with different aspects of cost-benefit analysis. The paper by Parsonson is concerned with developing a technique to empirically measure costs of urban traffic congestion in terms of vehicle operating speeds and value of travel time. The paper describes a methodology for developing such time and cost data.

The paper by McIntosh and Quarmby describes procedures for evaluating movement costs and benefits consequent to changes in network and transportation management policies and for estimating the generalized behavioral and resource cost functions.

The paper by Fleischer discusses the magnitude of the cost-benefit ratio and whether an item should be considered a benefit or a negative cost. The reader will see from the discussion that there is some difference of opinion on the subject.

The paper by Claffey deals with car fuel consumption as affected by snow and ice and has significant implications for cost-benefit analysis. Claffey states that fuel consumption is increased by 50 percent on 2 inches of snow as compared to dry pavement. This finding may have considerable impact on cost-benefit analysis of highway snow removal.

The paper by Yu and Wilhelm looks at the problem of how to optimize the capacity of transportation vehicle terminals. It is not economically feasible to design a terminal for the maximum demand only to have it substantially unused during the majority of the time. The paper develops a model for arriving at the optimum terminal capacity; the model takes into account the trade-offs, level of service, and cost effectiveness.

A SYSTEM TO MONITOR THE ROAD-USER COST OF URBAN TRAFFIC CONGESTION

Peter S. Parsonson, Georgia Institute of Technology

In any congested traffic stream, the road-user cost of time plus vehicle operation includes a component "cost of congestion" that is additional to the cost associated with flow at a minimum acceptable level of service of C. This project performed speed-and-delay runs using a tachograph-equipped car on Atlanta arterials and freeways. The resulting speed-time graphs were converted to dollar costs of congestion. The conversion was based on the tables of vehicle operating cost published in 1969 by Winfrey and on recent research on the value of time to automobile drivers and operators of commercial vehicles. The calculations were expedited by the computer program RUNCOST, written for this project by the Federal Highway Administration. Computer calculations of time cost plus operating cost were plotted against observed travel speeds. These plots yielded the congestion components of road-user costs. One hour of field data collection was found to require an expenditure of $15 for office processing. It was concluded that the monitoring system is both technically feasible and economical. Recommendations for congestion-monitoring programs and further research are presented.

•THIS PAPER is an abridged version of the final report of a recent research project on the road-user cost of urban traffic congestion. It was postulated, and demonstrated as part of the project, that congested flow is more costly to the driver than is flow at an acceptable level of service. There is an incremental "cost of congestion" that includes the dollar value of lost time plus any extra cost of operating a vehicle at an unacceptably low level of service. If it could be measured, the incremental cost of accidents would be another component of congestion cost. This project was directed toward techniques for measuring the total cost of time and vehicle operation in various traffic streams and for determining the component of that cost attributable to traffic congestion. Separating the congestion cost component from the portion that can be considered reasonable and acceptable makes it possible to obtain a true indication of the magnitude of the congestion problem.

NEED FOR CONGESTION-COST STUDIES

An economical method for measuring the dollar cost of congestion would be valuable for several reasons. One important use would be for mobilizing public and legislative support for proposed transportation improvements. These improvements need not be highway construction projects but could be proposals for public transit, traffic-signal improvements, or any other project aimed at reducing congestion.

If dollar costs of congestion were measured city-wide, and in several cities throughout a state, they would provide a basis for comparing the relative needs of these cities for transportation improvements. Congestion costs could be a basis for setting priorities

Sponsored by Committee on Transportation Information Systems and Data Requirements and Task Force on Application of Economic Analysis to Transportation Problems.

and would be of assistance in the difficult task of allocating state or federal transportation funds to the various urban areas on the basis of demonstrated relative costs of congestion. If city-wide measurements were carefully tabulated by travel corridor, the relative needs within a city would become apparent. The relative priority of a certain corridor project in one city versus a proposed corridor project in another city could be determined rationally.

It is emphasized that priorities would be determined by relative congestion cost, as defined earlier, rather than by relative total cost. The congestion cost reflects needs or deficiencies, whereas the total cost includes the portion considered reasonable and acceptable.

A third reason for congestion-cost studies pertains to the current Traffic Operations Programs to Increase Capacity and Safety (TOPICS). These programs involve operational improvements in signals, signs, markings, channelization, and the like to facilitate traffic flow without major new construction or right-of-way acquisition. In view of increased public resistance to new highway construction in urban areas, TOPICS projects are becoming increasingly important. Before-and-after studies of their effectiveness need to be sufficiently precise to reveal benefits on the order of 5 or 10 percent in some instances. TOPICS points up the need for a sensitive tool for precise before-and-after measurements of congestion costs.

Here again it is emphasized that attention should be directed to the congestion component of total road-user cost. The measure of effectiveness of a TOPICS project, for example, should not be the reduction in total road-user cost but rather the reduction in the congestion component. An improved signal system that has reduced total road-user cost by 12 percent may have reduced the congestion component by 90 percent!

REVIEW OF LITERATURE

There is no evidence in the literature of a technique for measuring the road-user cost of urban traffic congestion either for a selected roadway length or for an entire city. Consideration has been given to the road-user cost of stops, delays, and accidents in studies of traffic-signal systems, but typically these calculations are quite generalized, depend on average values, and do not take a microscopic look at the motion of an individual vehicle in the traffic stream. Moreover, they are concerned with the total cost rather than with the component attributable to congestion.

Past Practices

Measurements of urban traffic congestion have been made for many years by the well-documented methods of the speed-and-delay study or the travel-time study [National Committee on Urban Transportation (10), for example]. Although these procedures are based on a test vehicle "floated" in the traffic stream, they do not yield dollar costs, much less the congestion cost component. Instead, these procedures measure congestion as a delay rate, defined as the difference between the observed rate of motion and a rather arbitrary standard rate for that particular type of street.

Ten years ago there were insufficient data on vehicle operating costs. The major publication on the subject at that time (2) dealt primarily with passenger vehicles on rural roads. Data on the operating costs of trucks and buses were quite generalized. Operation at typical urban speeds, under stop-and-go conditions, was not well documented.

Until recently there was insufficient research on the value of lost time to automobile drivers and operators of commercial vehicles; many engineers felt that computations of monetary loss due to delay were controversial at best.

Some Recent Advances

Several recent developments have pointed the way to substantial improvements over past practices. A number of recording devices have been devised to aid in the gathering of data by a "floating" test car. Montroll and Potts (9) described and Argo-Kienzle tachograph that attaches to the speedometer cable of a test vehicle and furnishes a graph of vehicle speed. They used this device successfully in their research on acceleration noise.

Greissman (7) described his Traffic Data Compiler (available from Marbelite) of similar installation; it provides a speed graph and several digital readouts of speed-and-delay data.

Information on vehicle operating cost was greatly expanded and improved by Winfrey (12). Winfrey's comprehensive tabulations of vehicle operating costs for a wide range of uniform speeds and speed changes are appropriate for urban traffic composed of both cars and trucks. The time-speed charts furnished by the Argo or Marbelite devices give a complete account of a test vehicle's motion as it is floated through traffic. Therefore, they are ideally matched to the Winfrey tables for the calculation of operating costs.

Recent research has shed much light on the value of lost time to automobile drivers and operators of commercial vehicles. Thomas and Thompson (11) documented the value of time for commuting motorists as a function of their income level and amount of time saved. Adkins et al. (1) developed the values of time savings of commercial vehicles in various U.S. locations.

Sufficient research has also been performed on level-of-service criteria to allow the total road-user cost of operation and time to be divided into a component of cost associated with reasonable and acceptable traffic flow and a cost component attributable to congestion. The Highway Capacity Manual (8) set forth quantitative guidelines for acceptable flow in terms of minimum speed for freeways, signalized major arterials, and other types of roadways.

PURPOSE AND SCOPE OF PROJECT

This project has brought together these recent advances and has added computerized data processing for the purpose of producing an economical new system for measuring urban traffic congestion. The project tested the technical feasibility and the economy of routinely performing speed-and-delay runs with a tachograph-equipped car and then of converting the resulting speed graph to a dollar cost of congestion with the aid of a computerized version of Winfrey's tables of operating cost.

The testing program took place on selected sections of Atlanta freeways and arterials. Multiple speed-and-delay runs were performed using two types of commercially available tachographs. Manual classification counts were obtained concurrently to give the composition of the traffic by type of vehicle. These field data were then processed to yield road-user costs of congestion.

The final report for the project describes additional field data and office calculations that are not within the scope of this paper. These include supplementary machine volume counts, data on factors influencing capacity, and calculations of volume-capacity ratios. The final report also includes travel-speed contour maps prepared for typical sections of freeway and arterial roadways.

METHOD

Field Data Collection

Speed-and-delay runs were performed in 1970-71 on 13 selected sections of the Atlanta network with a total length of 85 miles. Six of these were Interstate freeway sections, each including several interchanges, and several were major arterial sections, each including a number of signalized, at-grade intersections. Two commercially available tachographs were tested; most of the runs were made with a Marbelite Traffic Data Compiler, but an Argo-Kienzle tachograph was used toward the end of the project.

The lengths of the sections varied from 3.7 to 10.4 miles, averaging 6.5 miles. The lengths were selected to be short enough to allow at least three speed-and-delay runs in each direction during the morning commuter rush and again during the afternoon peak. Three runs were also made during off-peak hours. All runs were made on a typical weekday, and complete data were recorded for both directions of the runs.

The Marbelite tachograph used for most of the speed-and-delay runs is a typewriter-sized device that rides on the front seat of the car alongside the driver. Priced at about \$3,000, it is driven by a connection to the vehicle's speedometer cable and is powered

by the vehicle's battery. It produces a continuous graph of the speed of the vehicle as it is driven through traffic. Unlike most commercial tachographs, however, the device includes several digital readouts of total trip time, total stopped time, and so forth. Photographs, sample speed charts, and field sheets were published by Greissman in 1967 (7).

Although it is possible for the driver alone to operate the Marbelite tachograph, especially in low-speed, stop-and-go traffic, an observer accompanied the driver on this project so that more detailed data could be recorded without compromising safety. During each run the observer made marks on the speed graph at a number of checkpoints. The marks were numbered consecutively, and the observer entered on a field sheet an identification of the location of each numbered checkpoint. The vehicle's odometer reading at each checkpoint was also entered and later became the basis for determining the length of each subsection.

Manual classification counts were taken concurrent with the runs in order to classify the composition of the traffic stream into the five typical vehicles for which Winfrey published data on operating costs. These counts were taken by a team of two observers stationed at a selected location in the section. One observer counted the passenger cars and the four-tired (light) trucks, while the other counted the three types of heavy trucks corresponding to Winfrey's cost tables. The team counted only one direction at a time. The counts were taken in one direction for a 5-min period, followed by a 1-min break for recording the tallies. Then counts resumed for 5 min in the other direction, followed by a 1-min break. In this way, the volumes in each direction were counted for 5 min in every 12-min period.

During the field work it was noted that the traffic conditions tended to vary widely within a section. Inasmuch as the sections are radials, the degree of congestion tended to decrease substantially with distance from the city center. Therefore, it was decided to divide the speed graphs from each section into subsections at intermediate checkpoints so that congested locations could be properly identified. The subsections were selected after consideration was given to the available volume-count records and the desirability of avoiding subsections so short that speed-and-delay results might be unstable. Figure 1, a sketch of an example freeway section, shows the five subsections into which the data from each run were subdivided and includes the ADT of each subsection as an indication of relative traffic use.

The Argo tachograph was also field tested. This compact unit mounts conveniently beneath the dash, is driven by a speedometer cable, and is powered by the vehicle's battery. The model used by Montroll and Potts (9) for traffic research purposes, model TCO-11/7G1K1, has two features vital for traffic engineering work. One is that the circular speed graph rotates once in 24 min (rather than 24 hours) and therefore can easily be read to the nearest 2 sec. A pack of seven 24-min charts permits continuous operation for up to 168 min. Also, this model includes an "event recorder," similar to the one on the Marbelite device, that records a mark on the graph at the push of a button. Inasmuch as the speed chart is inaccessible during the runs, the event recorder is quite necessary for locating checkpoints on the chart. The Argo unit is priced at approximately $300, including an "analyzer stand" that magnifies the small charts for easier reading in the office.

It is desirable to use an inexpensive dash-mounted clock and a battery-operated tape recorder with the Argo tachograph. At the beginning of the run and at each checkpoint the driver should actuate the event recorder momentarily, turn on the tape recorder, and record the time, the odometer reading, and a description of the event, such as "start of southbound run number two." The tape recorder is then turned off until the next checkpoint. The driver will have no difficulty in performing these functions, even on a high-speed facility. The final report of the project includes detailed instructions for the use of the Argo tachograph and its accessories.

Office Processing of Data

Computer Determination of Vehicle Operating Cost—After the speed graphs of the runs were divided into subsections, they were coded for computer calculation of vehicle cost.

These calculations were performed by the computer program RUNCOST (14), which was written in 1970 for this project by Bloom and later modified by Radics, both of the Federal Highway Administration. The purpose of the program is to eliminate the tedious process of translating a graph of speed versus time into a vehicle operating cost by means of Winfrey's tables.

The Winfrey tables give the costs per vehicle-mile of operating a passenger car and four types of trucks at uniform speeds ranging from 0 to 80 mph and also indicate the additional costs of accelerating or decelerating these vehicles. The tables take into account the profile gradients of the roadway and the horizontal curvature as well. Operating costs, more precisely termed running costs by Winfrey, include costs of fuel, tires, engine oil, maintenance, and depreciation. The cost of the fuel component does not include the state or federal road-user tax.

Briefly, RUNCOST uses the following six program control cards:

1. Title, which provides a heading that is printed at the top of each page of output, plus an adjustment factor (to correct for known tachograph error in recording speed) and a cost inflation factor (to update Winfrey's costs to the present);
2. GRAD, which provides information on the grade distribution;
3. CURV, which provides information on the horizontal curvature;
4. PAR, which provides the length of run in miles, the tachograph time scale calibration, the cost in dollars of an hour of vehicle time, and the total number of vehicles using the roadway during the time for which the speed-and-delay run is considered representative of traffic conditions;
5. VEH, which describes the distribution of the five vehicle types using the roadway; and
6. GO, which marks the end of a set of control cards and the beginning of the input data.

The input data cards describe the graphs of speed versus time as a series of coordinates. The coding of the graphs requires that they be digitized, that is, approximated by a series of points connected by straight lines. Each point has digital coordinates of time and speed that are coded on the input data cards.

The RUNCOST program prints out the following calculations of operating cost: cost per vehicle, tabulated by the five types of vehicles; cost per average (composite) vehicle; cost component due to speed changes and stops; cost component due to uniform speeds (on prevailing profile grades and horizontal curves); and cost per vehicle-mile of travel (VMT) for each vehicle type and for the composite vehicle. Additional printout includes stopped time, total travel time, the overall travel speed, and the length of the run as computed by the series of coordinates of time and speed. Computer printout of the cost of time is described next.

Computer Determination of Time Cost—Apart from Winfrey's tables, the RUNCOST program also computes the dollar value of the time of the run for each of the five vehicle types and for the composite vehicle based on dollar values of time specified by the user on the PAR card. The sum of operating cost and time cost for each type of vehicle is reported as well.

Thomas and Thompson (11) reported the value of time for commuting motorists as a function of their income level and the amount of time saved. In this project, therefore, a study was made of Atlanta income levels and travel characteristics so that the approximate value of time could be specified for passenger cars.

The average family income level for the five-county Atlanta metropolitan area was found to be in the \$10,000 to \$12,000 range as of 1968. The "amount of time saved" was more difficult to deal with inasmuch as speed-and-delay runs report a total time rather than a time saved by taking an improved or alternate route. Nevertheless, a hypothetical or typical amount of time saved was developed for Atlanta as follows. First, it was determined from the Atlanta Area Transportation Study (3) that the modal work trip length is 24 min. Next, Carter (13) found in his Wisconsin Avenue study that improvements on a signalized arterial can increase travel speeds from 20 to 30 mph. Such an increase in speed would mean a saving of 8 min in the modal work trip length in Atlanta. An 8-min saving was considered to be representative of other types of improvements also, for the purpose of this project.

Figure 1. Example freeway.

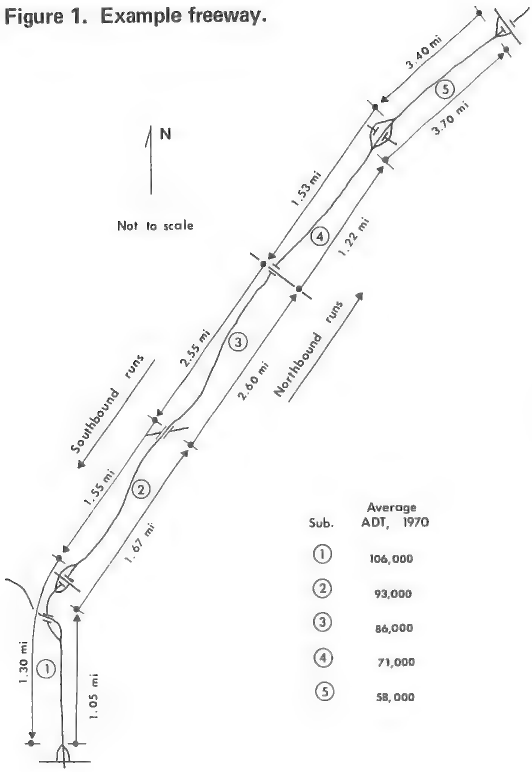


Figure 2. Cost versus speed for afternoon peak on example freeway.

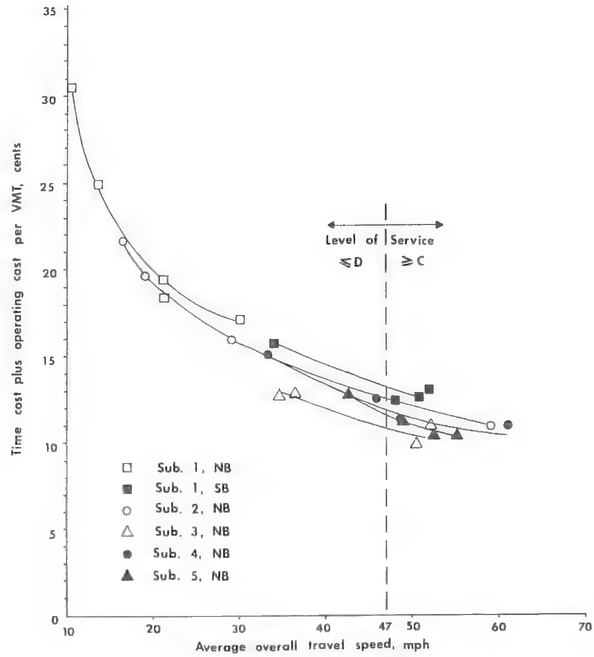


Table 1. Congestion costs for run 4, subsection 1, during peak period.

Cost Item	Cents
Operating cost/VMT	7.28
Time cost/VMT	23.15
Subtotal	30.43
Less estimated cost at 47 mph	12.0
Total	18.4

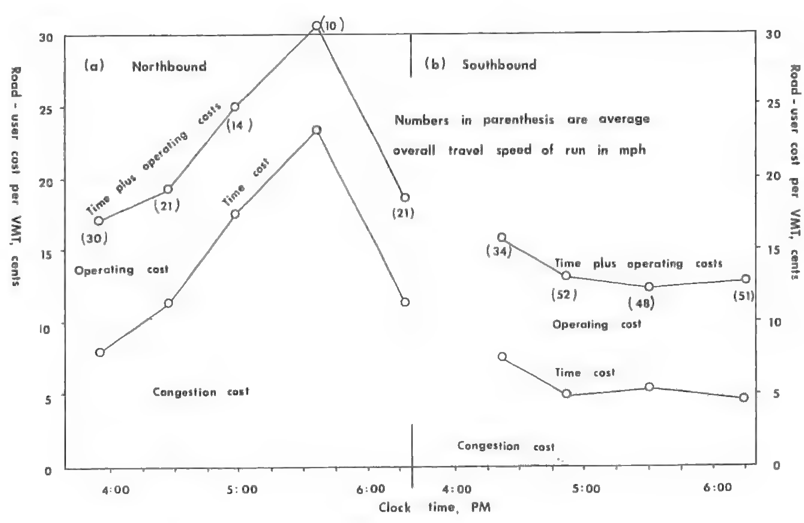
Note: Time of run was 5:36 p.m., and the average speed was 10.3 mph.

Table 2. Time and cost for office processing data.

Item	Time Expended (hour)	Cost per Hour (dollar)	Total (dollar)
Tape recorder playback	0.15	3.00	0.45
Processing of manual counts	0.15	3.00	0.45
Coding of control cards	0.15	3.00	0.45
Coding of data cards	2.0	3.00	6.00
Computer card punching	1.34	3.00	4.03
Computer cost (IBM 360/65)			1.20[a]
Portion of road-user cost due to congestion	1.0	3.0	3.00
Total			15.58

[a]See text for assumptions affecting this cost.

Figure 3. Congestion costs for subsection 1 of example freeway.



Thomas and Thompson indicate a value of time of $1.44 per person per hour for these levels of income and time saved (11, Table 5). With an average car occupancy of 1.5 persons, the value of time was calculated to be $2.16 per hour.

The values of time for Winfrey's four classes of commercial vehicles were obtained from Adkins et al. (1) and were updated from 1965 to 1970 on the basis of data from the U.S. Bureau of Labor Statistics and Dodge Trucks, Inc. The results were $4.23, $5.36, $6.26, and $7.06 per hour for trucks weighing 2.5, 6, 20, and 25 tons respectively (including driver's wages).

Obtaining Congestion Costs From Computer Output—The speed graphs and the computerized Winfrey tables were used to calculate road-user costs by subsection for all runs, whether or not congestion was present. The road-user cost attributable to congestion was determined by considering that congestion costs are accrued whenever the level of service falls below C, as defined by the Highway Capacity Manual (8).

In the case of freeways, the Manual considers the level of service to fall below C whenever the average speed is less than approximately 47 mph (corresponding to an operating speed of 50 mph) or the volume-capacity ratio exceeds 0.75. For urban and suburban arterials, the lower limit of level of service C is an average overall travel speed of 20 mph or a volume-capacity ratio of 0.80 to 0.90 (depending on the degree of signal progression).

To determine for each section the road-user costs attributable to congestion required that the level of service be determined on the basis of speed alone. The omission of the volume-capacity criterion simplified the determination of level of service and kept requirements for field data collection and office processing to a tolerable level for routine monitoring of congestion cost.

A graph was made of time cost plus operating cost per VMT versus average overall travel speed of each run. These data were obtained from the RUNCOST computer output. Figure 2 shows an example of this type of plot. The time cost plus operating cost corresponding to the lower limit of level of service C was then obtained from the graph. This lower limit is at 47 mph for freeways and 20 mph for arterials. Then, for any run, the portion of the time cost plus operating cost that is in excess of this graphical value was taken to be the congestion cost of the run. An example of this calculation is given in Table 1. Both Figure 2 and Table 1 are considered in greater detail later.

FINDINGS

Office Procedures

Table 2 gives the steps in office procedure that normally would be followed in routine measurements of congestion cost. The time and expense estimated for each step are also shown. Table 2 indicates that 1 hour of field speed-and-delay data will require office processing costing approximately $15. The largest single item of expense is seen to be the coding of data cards. The cost of computer processing per hour of field run will vary within wide limits, depending on such factors as size of batch processed, number of other users sharing the cost of the execution time, and the installation's policy on per-hour charges. In this project it was found that execution time ranged from 1 to 4 min per hour of field run; a time of 2.5 min is representative. Also, 46,000 bytes of storage are required for the execution of this program; therefore, nine other users could share the capacity and cost of the IBM 360/65. Assuming computer time to be valued at $280 per hour, the cost of computer processing was calculated as $2.5/60 \times \$280 \times 1/10 = \$1.17$ (rounded to $1.20).

Example Highway Subsection

The findings for subsection 1 of an example freeway (Fig. 1) during a single period of the day are presented herein. The manual counts taken during the afternoon peak indicated that, in the northbound (outbound) direction, the vehicle distribution was 90.4 percent passenger cars, 7.4 percent commercial delivery (2.5-ton) trucks, 1.4 percent six-tired, single-unit (6-ton) trucks, 0.5 percent semi-trailer, 20-ton trucks, and 0.3 percent semi-trailer, 25-ton trucks. In the southbound (inbound) direction, only 86.2

percent of the vehicles were passenger cars, and the truck percentages were correspondingly higher. For the vehicle distribution in the northbound direction, the weighted value of time was found to be $2.39 per hour and, for the southbound direction, $2.53.

Figure 2 shows the findings from the speed-and-delay runs on all five subsections with regard to average overall travel speed and road-user cost. The figure indicates that speed varied over a wide range, from 10 to 60 mph. The highest speeds tended to be associated with the lowest road-user cost—approximately 10 cents per VMT—whereas the lowest speed of 10.3 mph was associated with a road-user cost of over 30 cents per VMT. At a speed of 47 mph, which is the lower limit of level of service C, the road-user cost varied from approximately 10 to 13 cents per VMT. This variation is due to the fact that any particular average speed, such as 47 mph, can be associated with a wide range of road-user costs, depending on whether the vehicle maintains a uniform speed or experiences considerable speed changes ("acceleration noise"). A value of 12 cents, close to the average for all subsections, was taken as an estimate of the cost at 47 mph for subsection 1. This estimate of 12 cents was used in Table 1 to calculate the cost of congestion. Data given in Table 1 show that for an example run over subsection 1 the total road-user cost of operation and time was 30.4 cents. With the cost at 47 mph estimated to be 12.0 cents, the congestion cost for the run was found to be 18.4 cents.

The data from the calculations given in Table 1 are shown graphically in Figure 3. The figure shows for subsection 1 the average speed, operating cost, time cost, and congestion cost for each of the five northbound and four southbound runs by time of run. For example, at 5:36 p.m. in the northbound direction, Figure 3 shows plotted from Table 1 the values of time plus operating cost, time cost alone, and congestion cost. The figure indicates much greater congestion cost in the northbound (outbound) direction than in the southbound (inbound) direction, as might be expected during the afternoon commuter rush.

CONCLUSIONS

The specific objective of this project was to develop and evaluate a practical, economical, and rational system for monitoring urban traffic congestion and the associated road-user cost.

Based on the findings, it is concluded that the project was successful in demonstrating the feasibility of monitoring congestion and its cost. It was shown that the engineering profession now has available to it a new, precise tool for expressing the results of speed-and-delay runs in dollars-and-cents terms.

It is further concluded that the office procedure, based on the RUNCOST computer solution of Winfrey's cost tables, was found to be sufficiently economical in time and money to recommend itself for widespread use.

RECOMMENDATIONS

Program for Monitoring an Entire Urban Area

The following recommendations are offered for the development of a congestion-monitoring program in an urban area:

1. Travel corridors in the urban area should be identified.
2. Speed-and-delay runs with a suitably equipped test vehicle should be performed for each corridor annually during morning and afternoon peaks of a typical weekday and during any other important peaks caused by recreation, shopping, and so forth. The lengths of these runs should be selected so that consecutive runs will not be more than about 20 min apart.
3. If possible, these runs should be performed on the same day that the routine annual machine volume counts are scheduled, preferably recorded by hour and direction. Manual classification counts should be performed only if the existing file of such data is inadequate.
4. These runs may require supplemental delay studies at certain intersections. Speed-and-delay studies usually record through movements only and may not adequately reflect serious left-turn delays.

5. The office procedures of coding and computer processing should next be performed, as given in Table 2. This step includes the processing of the computer printouts to give the dollar cost of congestion per VMT, as in Table 1.

6. The next step should be a comparison of congestion costs per VMT with those measured in previous years to indicate trends with time.

7. Plots of congestion cost similar to those shown in Figure 3 should be prepared for selected sections, as needed for visual aids in describing a particular congestion problem.

8. Congestion costs per VMT should then be converted to congestion costs per mile by multiplying by the number of vehicles using the section.

9. Step 7 should be repeated, using congestion costs per mile, to indicate trends with time. Are these changes in line with the advance-planning forecasts of trends in traffic demand?

10. Again using congestion costs per mile, the sections should be ranked in order of congestion. Comparisons of corridor congestion can be made among the corridors within an urban area and among corridors in different cities. As relative needs become apparent, decisions concerning priorities and programming can be considered.

Further Research

1. The most urgent need for further research is in the area of accident costs of congestion. Accident costs per VMT could have been estimated satisfactorily for the 13 highway sections of this project. However, when the portion of road-user costs attributable to congestion are calculated, accident costs can be taken into account only if they are known as a function of level of service. Specifically, the relationship between accident cost per VMT and average travel speed, for various types of facilities, is needed.

2. A less expensive procedure is needed for digitizing the speed graphs for computer processing. A state-of-the-art review of this area is needed, followed by a determination of the most economical way to obtain access to the appropriate equipment.

3. It is to be expected that a large volume of road-user cost data will be generated by this monitoring system. These data should stimulate research along the lines of NCHRP Project 2-7 (6). The purpose of this project by Claffey (6; see also 5) was "to provide data on road-user costs as classified by arterial type, operating speed, traffic composition and delay factors." It is to be expected that the road-user cost data obtained by the monitoring system described herein will complement Claffey's work. In particular, these data will aid the development of basic tables applicable for planning and for selecting arterial street and expressway systems from the various alternates in urban areas. The Chicago Area Transportation Study (4) made use of such tables.

4. Many other research-oriented analyses of these road-user cost data suggest themselves immediately, such as the determination of the relationship between volume-capacity ratio and road-user cost (or congestion cost). Also, graphs of road-user cost versus average speed (such as Fig. 2) prepared for various types of roadways and traffic conditions need to be analyzed in order to determine a rationale for curve shape and location.

ACKNOWLEDGMENT

This work was performed at the Georgia Institute of Technology under contract with the Georgia State Highway Department, in cooperation with the Federal Highway Administration. The opinions, findings, and conclusions in this paper are those of the author and not necessarily those of the state highway department or the Federal Highway Administration.

The author is grateful to the Federal Highway Administration for the development of the RUNCOST computer program for this project and for making available the FHWA computer facility and personnel for all of the computer runs. The author is also indebted to the Georgia Highway Department for performing all of the computer card punching for this project.

REFERENCES

1. Adkins, W. G., Ward, A. W., and McFarland, W. F. Values of Time Savings of Commercial Vehicles. NCHRP Rept. 33, 1967.
2. Road-User Benefit Analyses for Highway Improvements. AASHO, 1960.
3. Atlanta Area Transportation Study. A Report on the Development of a Trip Distribution Model. Div. of Highway Planning, Georgia State Highway Dept., 1968.
4. Chicago Area Transportation Study, Vol. 3, Transportation Plan. Chaps. II and III, 1962.
5. Claffey, P. J. Running Cost of Motor Vehicles as Affected by Highway Design—Interim Report. NCHRP Rept. 13, 1965.
6. Claffey, P. J. Running Costs of Motor Vehicles as Affected by Road Design and Traffic. NCHRP Rept. 111, 1971.
7. Greissman, J. The Traffic Data Compiler. Public Safety Systems, May/June 1967.
8. Highway Capacity Manual—1965. HRB Spec. Rept. 87, 1965.
9. Montroll, E. W., and Potts, R. B. Car Following and Acceleration Noise. HRB Spec. Rept. 79, 1964, pp. 37-48.
10. Determining Travel Time. National Committee on Urban Transportation, Public Administration Service, Chicago, Procedure Manual 3B, 1958.
11. Thomas, T. C., and Thompson, G. I. The Value of Time for Commuting Motorists as a Function of Their Income Level and Amount of Time Saved. Highway Research Record 314, 1970, pp. 1-15.
12. Winfrey, R. Economic Analysis for Highways. Internat. Textbook Co., Scranton, Penn., 1969.
13. Carter, A. A., Jr. Increasing the Traffic-Carrying Capability of Urban Arterial Streets—The Wisconsin Avenue Study. Div. of Traffic Operations Research, Office of Research, Bureau of Public Roads, U.S. Dept. of Commerce, 1962.
14. User's Guide to RUNCOST. Traffic Eng. and Parking Branch, Urban Planning Div., Federal Highway Administration.

GENERALIZED COSTS AND THE ESTIMATION OF MOVEMENT COSTS AND BENEFITS IN TRANSPORT PLANNING

P. T. McIntosh, Strategic Planning Directorate, Department of the Environment, London; and
D. A. Quarmby, London Transport Executive

The object of the paper is to provide guidance to transport planners and analysts by describing procedures in two areas: (a) the evaluation of movement costs and benefits consequent to changes in networks and management policies and (b) the estimation of the generalized behavioral and resource cost functions for links and origin-destination pairs that are necessary for this evaluation process and for forecasts of behavior. The procedures are designed for use in situations where the change in network or policy is thought to have strong effects on the trip pattern and individual link loadings. This will generally be the case in the consideration of urban schemes and may be the case for major interurban schemes; in both situations there may be considerable changes in the trip matrices, modal split, and routes used. The emphasis is on operational methods. The precise way in which the benefit expression and generalized costs are calculated will depend on the level of detail and form of particular studies; considerable guidance is given to aid the transfer from concepts to computation.

•TRAFFIC PREDICTION was, for many years, carried out quite independently of the procedures used for assessing the economic value of the possible changes under consideration. In the late 1950s and 1960s, the principal outputs of traffic models were flows of people and vehicles along links of networks in urban areas, and investments were very largely decided on by considerations of physical and technical feasibility (operational evaluation). At the same time, techniques were evolved for estimating the movement costs and benefits arising from the improvement of particular roads, mainly in rural areas; in such situations, the facilities of traffic models—the ability to represent the response of traffic movements over a wide area to changes in the road network—were considered unnecessary in the evaluation procedures.

A growing desire within the responsible authorities not only to obtain value for money in transport investment but also to make comparisons between feasible options has led in recent years to a strong need to integrate the methodology of traffic models with that of economic assessment procedures. The London Transportation Study was probably the first to attempt to do this in network comparisons (1, 2, 3), and a procedure for isolated road schemes was described by the Road Research Laboratory in 1968 (4). The problems were, and largely still are, substantial: To start with, the languages of traffic prediction and of highway investment appraisal were fundamentally different; traffic prediction methods evolved empirically, as a collection of heterogeneously based submodels, with no explicit economic inputs and no economic basis at all; as applied, highway investment methods were aimed at the consideration of individual links or small schemes and, based almost entirely on a "travel time and cost-saving" approach, were unable to handle, other than very simply, consumers' surplus aspects arising from changes in traffic behavior.

Sponsored by Task Force on Application of Economic Analysis to Transportation Problems.

The establishment of a common basis and framework for both traffic prediction and economic evaluation has developed along a broad front within the Department of the Environment (formerly the British Ministry of Transport) and elsewhere as a result of work by both mathematical analysts and economists. The associated methods are now beginning to be used as a matter of routine in planning projects carried out within the Department. The intention of this paper is to disseminate both the thinking and the methodology among local and regional government authorities, public transport authorities, consultants, and others engaged in transport planning so as to establish consistency between studies and to facilitate better predictions and an improved allocation of investment resources.

The layout of the paper reflects the two main themes of this integrated approach to prediction and evaluation. The first part deals with the introduction of an explicity economic content into characterization of space and time in traffic prediction, by means of the concept of "generalized cost"; the nature of the cost function is related first to the factors that influence travel behavior, and then to the consumption of real resources that come about through changes in this behavior. Then there is a section concerned with broadening the cost-saving approach to economic evaluation so as to include the measurement of changes in consumers' surplus that arise through changes in travel behavior as predicted by the traffic models.

THE CONCEPT OF GENERALIZED COST

For some time transportation analysts have been aware that travel time alone is not a satisfactory way of representing the separation between zones as used in transportation studies, particularly for modeling people's travel behavior. For one-mode forecasting (e.g., road) time may be an adequate measure, although the inclusion of high-speed roads in test networks can highlight the problem, inasmuch as the extra mileage and operating costs in a given travel time at a high speed remain hidden. To some extent this effect has been concealed in the general "noise" implied in the present levels of accuracy attained by traffic models.

However, modeling more than one travel mode with an integrated model (5) (as opposed to ad hoc techniques such as distributing trips with road "skim trees" and then using diversion curves) exposes the limitations of time alone rather more obviously, principally because time-cost profiles can differ significantly between travel modes. Furthermore, time-based models are insensitive to changes in the pricing of public transport and of parking facilities, and they can say nothing about road pricing. Operationally, the analyst can express, for example, parking costs as so much extra time on a link, but this is only an ad hoc method of dealing with a problem that it is more satisfactory to tackle in a basic, more general way.

Because the use of integrated distribution-modal split models is increasing, the need is clear to establish a framework for the definition of a more generalized measure of the costs of travel as represented in model-based studies to replace time in the specification of networks for use in distribution, modal split, and assignment procedures.

In addition, the definitions of cost developed to represent how people's travel behavior depends on the characteristics of networks are also relevant in putting an economic interpretation on their behavior. The sections on the estimation procedure show how this "behavioral cost," used to explain travel demand, is also used to attribute "value" received by travelers when networks are improved or policies changed.

At the same time, the economic evaluation procedure is concerned with estimating the real resources consumed in travel and transport. A further concept, parallel to behavioral cost, is needed—a "resource cost." This is a unit cost that describes the value of resources consumed in a unit of travel.

Why should these two costs be different? A behavioral cost is that cost function that best explains people's travel behavior (and therefore enables their behavior to be predicted). A resource cost is that cost function that represents a consumption of resources. Thus the following two areas can give rise to difference:

1. People may base their behavior on imperfect perceptions of cost. For example, there is considerable evidence to suggest that people significantly underestimate the

costs of running cars (6, 7, 8). The mileage cost that best explains and models their travel behavior is less than a strict engineering assessment of the marginal mileage cost.

2. The prices that people face (and therefore that determine their behavior) may not reflect the true resource costs. For instance, taxation on fuel is a component of the price but does not represent resources; also, the fares charged on a public transit system may not reflect the actual costs of operating the different parts of the system, and so on.

Behavioral costs for prediction will normally be measured in equivalent time units. In the evaluation procedure where value received is calculated, a special form of the behavioral costs is needed. This differs in two ways from the prediction value: (a) it is calculated in common monetary units (simply a change of scale), and (b) it may contain assumptions on the values to be attributed by the community to certain items that are different from those placed on those items by the individuals themselves (e.g., a commor value of nonwork time rather than a behaviorally revealed value differing according to income group). In this way the effect of weighting benefits in various ways relatively between groups of recipients can be examined without affecting the forecasting procedures.

Thus for the prediction of travel demand and the appraisal of transport investments and management policies there are the following three kinds of generalized costs:

1. b, behavioral cost for use in prediction models; form of the function is based on the best knowledge about what characteristics of networks influence people's and firms' travel and transport decisions. It takes into account time and costs and is usually in time-equivalent units.
2. u, behavioral costs for use in the benefit estimation procedure; in current practice the form of the function is identical to b except that it may include alternative values of nonwork time as reflections of possible social values and that it is in monetary units.
3. r, resource cost for use in the benefit estimation procedure as society's valuation of the resources consumed by a unit of travel; the form of the function is based on known technical relationships between costs and various transport-related activities. Some items in the function are based on behavioral cost items; they will use the values from u, not b, where they differ.

Theoretical Aspects of the Derivation of Behavioral Cost Functions

Behavioral cost is an expression describing the totality of "cost" or disutility incurred by a traveler in making a zone-to-zone trip by a particular mode of travel; it may well not be the total cost or disutility that the traveler actually incurs, because he may have an imperfect perception of cost. In practice it is simply the cost that best explains his travel behavior within the framework of the model processes in use. And, inasmuch as traffic models are trying to represent people's travel behavior, this is the right sort of cost to use for prediction.

Operationally, all the factors contributing to travel disutility are not known, nor could they all be included in present modeling procedures. In addition, each traveler will implicitly behave according to a unique set of factors (e.g., traveling time, waiting time, fares, interchanges, and comfort) and a unique relative weighting of them. The models used are concerned with the behavior of travelers in aggregate; thus, a decision must be made on the level of generality of the behavioral cost function to be used. Ideally, both the form of the function and the values of its parameters should be relevant to the particular study and possibly determined from data gathered in the area where the transport model is being applied. The main arguments against this are first that most transportation studies cannot sensibly mount the program of survey and analytical work that would be necessary; and second that, with the uncertainties that must invariably attach to estimates of values from individual research studies of this kind, more robust parameter values and function forms can be arrived at by joint consideration of several research studies, providing that the results are expressed in a sufficiently generalized form.

The object therefore is to define a cost function and to suggest parameter values for use in studies. The form of cost function suggested here is

$$b_l = B_1 x_1 + B_2 x_2 + \ldots + B_n x_n$$

where b_l is the behavioral cost of travel along a link l of a network by a particular travel mode. $x_1, x_2, \ldots, x_n$ are values of factors that are important in determining the overall travel disutility as it affects behavior. $B_1, B_2, \ldots, B_n$ are the relative weights of these factors. More complex functions could be suggested, but at this time there is no good reason for not using a simple linear function, especially as it has additive properties that are plausible and simplify calculation.

Normal network manipulation and tree-building programs can then be used to find cheapest routes (using behavioral cost), instead of fastest routes. Trees are skimmed to produce an interzonal behavioral cost matrix, b_{ij}, instead of an interzonal travel time matrix. (If more sophisticated multiple-route finding procedures are used, similar arguments apply.) The behavioral cost matrix is then available for use in distribution, modal split assignment, and evaluation procedures.

In deciding what factors should be taken into account, one can start by including time and add other factors that seem reasonable. Alternatively, one can draw on research in related fields to discover what factors people seem, by an analysis of their behavior, to take into account in their travel. In the last few years, information has come to light from studies of people's choice of travel mode in the journey to work on the factors influencing that choice, and their relative weights (7, 8, 9, 10, 11). In that these factors seem to be taken account of when people compare one mode of travel with another, it is sensible to impute that this is how people see each mode individually.

As a result of this and other empirical work, it is recommended that the factors that should be included in a behavioral cost function are the in-vehicle travel time, the components of excess or outside-vehicle travel time (suggested split is walking time and waiting and transfer time), and the financial cost of travel (including terminal cost, if any). The behavioral cost function for a network link is thus

$$\begin{aligned} b_l = {} & B_1 \times \text{in-vehicle time} + B_2 \times \text{walking time} \\ & + B_3 \times \text{waiting and transfer time} + B_4 \times \text{travel cost} \end{aligned}$$

For any given network link, only some of these variables will be nonzero. For instance, a highway link in a private car network will use only the in-vehicle time and the travel cost (as mileage × cost per mile related to the speed). A terminal link may have a walking time and a parking charge. On a public transit network, an access link may contain a walking time and a waiting time, a route link will contain an in-vehicle time and a fare, and so on. The values of the various coefficients may well vary according to the trip purpose and income group under consideration. This particularly applies to the time values.

It can be seen that B_1/B_4 is the value in travel cost units of in-vehicle traveling time, B_2/B_4 the value of walking time, and B_3/B_4 the value of waiting and transfer time. For any particular study, therefore, the task is to estimate the relative values of the parameters, B, and to decide what units to express them in. It might seem obvious to express b_l in monetary units, but it will be seen in a subsequent section that it may be more appropriate to express the behavioral costs in time-equivalent units when forecasting behavior.

The behavioral costs for evaluation, u_l, will use the same function except, where appropriate, it will incorporate any alternative values under consideration—for example, where the effect of alternative values of nonwork time are being examined. They will be expressed in monetary units.

Resource Costs

A resource cost function is similar in form to a behavioral cost function; it contains personal time, valued at the appropriate rate, and engineering assessments of vehicle

or system marginal operating costs, less transfer and other nonresource payments such as taxation. For highway links, the engineering assessments of motor vehicle operating cost can be related fairly easily to the link and thus become part of link unit resource cost.

However, where the prediction units are person trips (rather than vehicle trips) and the vehicle occupancy is not necessarily constant (for instance, with bus operation), it may be unsatisfactory to express the resource costs as so much per person traveling. For fixed track systems, where the marginal and average costs diverge substantially, a unit rate per person trip is nearly always unsatisfactory or impossible to determine. In these cases, therefore, the unit link resource cost will only include personal time, and the change in system operating costs will be estimated separately on the basis of total network assignments or other system characteristics. The resource cost function will then be

$$r_l = R_1 \times \text{in-vehicle time} + R_2 \times \text{walking time} + R_3 \times \text{waiting and transfer time} + R_4 \times \text{unit resource costs of vehicle operation}$$

Cost Functions in Network Manipulation

The functions are specified as link costs. In this form they can be used to derive zone-to-zone (i-j) values for use in predictive models and in the evaluation procedure. The process is as follows:

1. Estimate link behavioral costs, b_l;
2. Build "cheapest" trees by normal network manipulation and tree-building procedures;
3. Skim these trees to provide zone-to-zone behavioral costs, b_{ij};
4. Using link u-costs, u_l, and link resource costs, r_l, add up each along the behavioral cost paths to obtain zone-to-zone u_{ij} and r_{ij}; and
5. Then the b_{ij} are available for the prediction model, and u_{ij} and r_{ij} are ready for the evaluation procedure.

There is an alternative approach: Some network computer programs only carry link stores for time, t_l, distance, d_l, and possibly speed, with no easy means of subdivision by classes h of time or special treatment for parking, fares, and the like. In this situation it may be useful to convert all the various time elements, t_{lh}, to equivalent "in-mode" time (scale by B_h/B_1) and put these in the time field; similarly cost or distance elements can be treated together and put, to some scale, in the distance record. Then b_l can be formed as needed in the program as a function of t_l and d_l, and b_{ij} can be found as before by building b-trees; t_{ij} and d_{ij} can be found by adding up times and distances along b-cost path, and then, in some circumstances, u_{ij} and r_{ij} may be formed from an appropriate linear function of t_{ij} and d_{ij}.

Difficulties will arise from the treatment of speed in r-costs and with components of the t_l and d_l that have different weights within b, u, and r; e.g., parking charges should appear in full in b but only the resource component should appear in r.

ESTIMATING BEHAVIORAL AND RESOURCE COST FUNCTIONS

Parameters Needed and Classification of Trips

In any particular transportation study, the values of the parameters B, U, and R in the following functions must be estimated:

$$b_l = B_1 \times \text{in-vehicle time} + B_2 \times \text{walking time} + B_3 \times \text{waiting and transfer time} + B_4 \times \text{travel cost (including terminal and toll charges)}$$

$$u_l = U_1 \times \text{in-vehicle time} + U_2 \times \text{walking time} + U_3 \times \text{waiting and transfer time} + U_4 \times \text{travel cost (including terminal and toll charges)}$$

$$r_l = R_1 \times \text{in-vehicle time} + R_2 \times \text{walking time} + R_3 \times \text{waiting and transfer time} + R_4 \times \text{unit operating cost}$$

The values of the parameters in fact depend on a large number of quite specific characteristics of the transportation study and of the models used, such as the units of

prediction (e.g., person trips or vehicles?), the base year, the prediction years, where modal split comes into the modeling process, and degree of disaggregation by purpose, person type, and income, and it is therefore not possible to formulate universal sets of standard values. Values of time, particularly of nonwork time, are notoriously difficult to estimate. However, they seem to have some generalized characteristics that help the estimation of local values for particular studies. A report of a British conference on the value of time (16) gives a reasonably up-to-date review of recent empirical and theoretical studies. As a general guide:

> In large urban studies it is usual to build models at one or more future years for person trips and to convert to private car or bus or train movements by applying occupancy factors after the modal split procedure. In some studies peak hour flows only are modelled, and in these the journey to work predominates. Commercial vehicles are ignored or treated separately outside the main modelling process. With this specification, the units would be person-trips, and the purpose-mix (which leads to value of time and occupancy) would reflect the peak period composition. . . . In inter-urban highway studies, it has been conventional, because of lack of data, to model vehicle trips, and to build in assumptions about occupancy and modal choice at the beginning, often implicitly. Often all vehicles (commercial, private cars and buses) are modelled together, so the unit cost would reflect an average purpose mix for the private cars and buses, and an average vehicle mix for the total traffic flows.

Further subclassification of the units of prediction may well be made in the more sophisticated studies, and it is a matter of choice for the analyst whether different cost functions should be developed for each subclassification. For instance, large urban studies may classify journey purposes into home-based work (HBW), home-based other (HBO), and non-home-based (NHB). The population may be stratified into different groups for the purpose of predicting travel behavior. One such classification is into car-owning and non-car-owning households. Another may be by income groups. In both these cases, different values of time will apply because of the different mean incomes. This, in turn, will alter the relative weighting of time and cost in the unit cost functions.

Ideally, therefore, the analyst should use a different cost function for each income group and purpose, which means different costs for any one link in a network and thus different networks. The extent to which this should be done will depend on the particular study and on the analyst's judgment. In many cases such stratification may not be necessary, but there may be circumstances where it is important to represent the fact that networks can really look different to people of different incomes (for instance, long-distance rail commuting to central London where financial costs of travel are relatively high compared with the travel time). In particular, it may often be that different networks should be built at least for car owners and non-car owners, using values of time based on incomes in each category.

Base Year and Rates of Growth

The values of all economic inputs, such as costs and time values, depend on the years in which modeling is attempted—the survey year for calibration and one or more forecasting years—and on the choice of a base year for prices. It is necessary to carry out all economic comparisons at some constant price level. It is conventional to assume that vehicle operating costs (both behavioral and resource) will remain constant at constant prices and that average values of time will rise at some assumed rate of growth of real incomes.

Where some disaggregation into categories of different incomes is adopted, separate rates of income growth should be estimated for each category. One curious effect of this is that the mean incomes of car-owning households and non-car-owning househoulds may both rise at less than the average rate. This comes about because of the acquisition of cars by non-car owners whose incomes tend to be high relative to other non-car owners and low compared with car owners; they "dilute" the car owners as car ownership increases.

Choice of Scales for the Unit Costs

At first sight it would seem obvious to scale the parameters so that the behavioral costs b are in money units. However, projecting values to a future date for forecasting travel patterns exposes a problem of consistency and comparability. As incomes rise, a given cost will carry less weight; it may be better to scale the parameters so that the units of behavioral cost retain some absolute value over time. It can be argued that time has much the same value in terms of personal utility to people of different incomes and to people living now and at some future date. There are, of course, arguments against this proposition, but at least it is probably more tenable than scaling on cost.

At the present time, therefore, it is recommended that, in forecasting procedures, the behavioral cost functions b be scaled on time, so the units are "equivalent minutes." This means that, for groups of higher income, the value of the cost component in the function will fall; i. e., a particular financial cost means less to those with higher incomes. For instance, if the real income (relative to prices in general) of some category is expected to rise by 50 percent (i. e., 1.5 times), then the coefficients of cost in the b functions fall by $\frac{1}{3}$ (i. e., $\frac{1}{1.5}$ times).

Effectively, as people become relatively better off, time assumes a greater proportion of the behavioral cost b of a trip. (For a typical trip to work by car in U. K. urban areas in 1968, time accounted for just about half the behavioral cost.)

In the evaluation procedure, however, resource costs, r, and modified behavioral costs used, u, should be in common monetary units.

Summary of Behavioral and Resource Costs

As a general rule, therefore, for each group considered, the following holds true:

1. In b-costs: B_1 will be unity, B_2 will be some factor such as 2, B_3 will be some factor such as 2, and B_4 will be one divided by the (averaged) value of traveling time, at the appropriate unit base (persons or vehicles).
2. In u-costs: U_1 will be the appropriate value of traveling time in monetary units, U_2 will be some factor such as $2U_1$, U_3 will be some factor such as $2U_1$, and U_4 will be unity.
3. In r-costs: R_1 will be the same as U_1, plus time-dependent elements of vehicle operating cost, R_2 will be the same as U_2, R_3 will be the same as U_3, and R_4 will be unity.

The factors for B_2, B_3, U_2, and U_3 are current estimates based on behavioral studies. If other locally derived or more reliable factors are available, they could be substituted.

ESTIMATION PROCEDURE FOR MOVEMENT COSTS AND BENEFITS

In transportation planning, economic evaluation involves the estimation of the costs and benefits that accrue as a result of investment in networks or changes in their management; its purpose is to make objective statements relating to the relative and possibly the absolute worth of alternatives. The procedures described here are designed for use in the situations where the change in network or policy is thought to have strong effects on the trip pattern and individual link loadings. This will generally be the case in the consideration of urban schemes and may be the case for major interurban schemes as well; in both situations, there may be considerable changes in the trip matrices, modal split, and routes used.

Individual networks or management policies cannot be sensibly considered in isolation. Comparisons between alternatives are essential for the valid consideration of the individual proposals. One particularly important comparison is between the various possible future systems under consideration and, as a base, the "do-nothing" situation; in this context "do nothing" means including only those changes to the existing situation that are, for all practical purposes, now unavoidable between the present day and the period under consideration.

The principal items giving rise to costs and benefits are

1. Capital and initial costs: (a) construction costs, including interchange facilities, parking provision, etc.; (b) land and property costs; and (c) delays and inconvenience during construction; and
2. Recurring costs and benefits: (a) transport user benefits and costs; (b) operator revenues and costs; (c) user transfer payments (e.g., taxation); (d) change in accidents; and (e) external economies and diseconomies (development consequences, environment, etc.).

This section concentrates on the measurement of those costs and benefits that arise from changes in the volumes and pattern of movement as a result of changes in networks or management policies. In particular, it deals with the estimation of transport user costs and benefits (item 2a) and their joint treatment with operators's costs and benefits (item 2b) and taxation (item 2c). Neither the estimation and treatment of the other important costs and benefits nor the problem of intergrating all these into a decision-making framework is considered here.

Why Not Simply Compare User Costs?

Many studies have calculated the relative benefits of alternative plans by simply calculating the change in user costs, on the basis of either a fixed or changed trip-making pattern. The following trivial example shows some dangers of this procedure.

Two towns, A and B, each have 10 internal or intrazonal trips taking place at 5 cost units each. The towns are connected by a poor road. It would cost a traveler 30 cost units to journey between them and, as a consequence, only one does in the time period considered.

Situation 1

	A		B	
Trips	10	1	10	
Cost per trip	5	30	5	Total cost 130

The connecting road is improved so that the cost is 15 units. Assume that the trip pattern is unchanged.

Situation 2

	A		B	
Trips	10	1	10	
Cost per trip	5	15	5	Total cost 115

Let the trip pattern and intrazonal cost now change in a sensible way.

Situation 3

	A		B	
Trips	8	6	8	
Cost per trip	4	15	4	Total cost 154

The total user cost (situation 2) with a fixed trip pattern is reduced below cost (situation 1) by 15 units. The cost (situation 2) with a sensibly changed trip pattern (as would happen in real life and in a transportation model sensitive to costs) is 24 units more expensive. A minimum cost comparison would then suggest that the improvement to the link was worth 15 units (situation 2) or else was worse than useless (situation 3).

Both solutions cannot be correct. The first solution shows a benefit but takes no account of the resulting change in trip-making; the second allows for the change in trip-making but indicates that the travelers as a whole are worse off.

The procedures described in the following overcome this dilemma by taking into account the benefits stemming from improved choice and the change in trip-making.

Consumer Surplus Approach to the Measurement of User Benefit

In general economic theory, this problem is well known and understood (12, 13). Only in recent years has the methodology been developed to evaluate the benefits in comprehensive comparisons of alternative networks (1, 2, 14, 15).

The measure of user benefits is essentially a measure of change in consumers' surplus. To estimate the benefit of a single plan and hence, by difference, the change in benefit between two plans, we should consider a demand curve for travel. This curve relates the demand q between a particular origin and destination to the behavioral cost of travel between them, b, defined here in monetary units; it is shown in Figure 1.

The definition of the curve indicates that for any particular trip at , say, q' the traveler would have been willing to pay b', thus making a profit or surplus of $b' - b$. When all these surpluses are added, they form the rather ill-defined upper shaded area in Figure 1, which is the consumers' surplus.

Some part of the cost to the users as indicated by their behavior b does not reflect consumption of resources (e.g., taxation or parking charges that may not equal the costs of provision and operation) but is additional surplus transferred either to the community as a whole (through taxation) or to operators (e.g., a parking authority). In Figure 1, then, the nonresource n element of the surplus is the lower shaded area, and the total surplus is the complete shaded area.

Figure 2 shows the benefit arising when two plans are compared. Suppose initially that the cost of travel is b_1, at which cost the demand for travel is q_1, but that as the result of a transport improvement or a different plan the cost of travel falls to b_2 and the demand for travel increases to q_2. The transport users obtain an increased consumers' surplus, illustrated by the upper shaded area in Figure 2, which is approximately $\frac{1}{2}(q_1 + q_2)(b_1 - b_2)$. In addition, there is the nonresource correction ($n_2 q_2$ -

Figure 1. The estimation of movement benefits: demand curve and surplus.

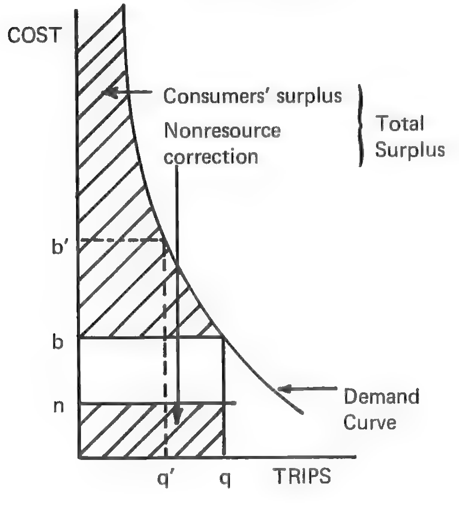
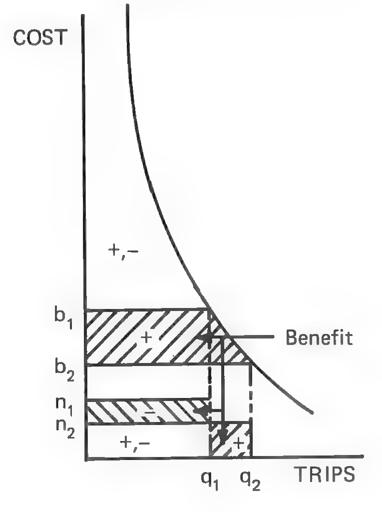


Figure 2. The estimation of movement benefits: comparison of alternatives.



$n_1 q_1)$, giving a total change in surplus or benefit of

$$\Delta = \tfrac{1}{2} (q_1 + q_2) (b_1 - b_2) + (n_2 q_2 - n_1 q_1)$$

Various algebraic manipulations that improve understanding and ease computation are possible. In particular we have $n = b - r$, where r is the unit resource cost, and hence

$$\Delta = \tfrac{1}{2} (q_1 + q_2) (b_1 - b_2) + q_2 (b_2 - r_2) - q_1 (b_1 - r_1)$$

or

$$\Delta = \underset{(A)}{\tfrac{1}{2} (q_1 + q_2) (b_1 - b_2)} + \underset{(B)}{[(q_2 b_2) - (q_1 b_1)]} - \underset{(C)}{[(q_2 r_2) - (q_1 r_1)]}$$

In other words,

Δ = increase in user surplus (A) ⎫ i.e., increase in gross value received by
+ increase in costs to users (B) ⎭ travelers as measured by behavioral costs
- increase in resource costs (C)

In this form it will be seen that the change in resource costs (C) can either be calculated from unit costs per trip or be estimated separately from a consideration of overall costs. For public transit systems, where there are many shared but variable costs, the latter course may be necessary.

This important result is the basis of the current network evaluation procedures, and in practice it is estimated for all the trip classes pertinent to the study; i.e., the expressions are summed over origin-destination pairs, modes, times of day and year, trip purposes, and person type classifications (e.g., income groups or car owners and non-car owners) to give the total direct movement costs and benefits. It is possible to examine partial summations (e.g., separate out origins, destinations, or person types) with a view to learning about the distribution of benefit. The validity and uses of this procedure are still under examination.

Strictly speaking, it is not possible to describe the demand curve for any one of these separate trip classes unless the costs for all the others are kept constant; to this extent the explanation and the figures are simplifications because the traffic models simulate the fact that the costs of travel between many origin-destination pairs will vary simultaneously when networks are altered. However, it can be demonstrated (15) that this treatment is a close approximation to the much more complicated situation where all the demand curves vary together. This simple expression also requires that the land use and socioeconomic assumptions be constant between alternatives.

Application of the Approach to the Example

Returning to the example, assume for simplification that the user costs quoted do in fact equal resource costs. In this situation there is only the increase in user surplus analogous to the attempted calculations in the example. Comparing the before situation with the after situation, which takes account of the changed trip-making and pattern, gives the user benefit of

$$\begin{aligned}\Delta &= \overset{AA}{\tfrac{1}{2} (10 + 8) (5 - 4)} + \overset{AB}{\tfrac{1}{2} (1 + 6) (30 - 15)} + \overset{BB}{\tfrac{1}{2} (10 + 8) (5 - 4)} \\ &= \tfrac{1}{2} (18 + 105 + 18) \\ &= 70.5 \text{ units}\end{aligned}$$

Note that this differs from both of the previous estimates. Previously there was a benefit of 15 units with the fixed trip pattern and a disbenefit of 24 units with the changed pattern. As explained, the difference is due to placing a value on the benefits that stem from the changes.

Behavioral Costs for Evaluation

The previous sections, for ease of presentation, derived the basic evaluation expression in terms of the behavioral costs, b. However, as explained earlier, in the

evaluation procedure the behavioral costs need to be in monetary units, and it may be appropriate to study the effect of values of certain elements differing from individuals' values as revealed through their behavior. This allows, for example, the effect of alternative relative weightings of benefit among time, cost, and income groups to be examined and different views on their absolute worth to be considered.

It is necessary then to calculate and use the variable u, which is identical in form to b except that it contains the alternative values and is scaled in cost units. The resource cost r is only defined in the context of the evaluation and will take account of any alternative values used in u. The expressions for the benefit become

$$\Delta = \tfrac{1}{2} (q_1 + q_2) (u_1 - u_2) + q_2 (u_2 - r_2) - q_1 (u_1 - r_1)$$

and

$$\Delta = \tfrac{1}{2} (q_1 + q_2) (u_1 - u_2) + [(q_2 u_2) - (q_1 u_1)] - [(q_2 r_2) - (q_1 r_1)]$$

Treatment of Taxation

As indicated earlier, some part of the difference between resource costs and behavioral costs is in terms of indirect taxation, which will account for part of the benefit. This will be transferred through the taxation system to society at large and arises because such taxation payments do not constitute use of resources; this nonresource element in the transport sector has already been taken into account. However, any increase in expenditure on transport would be accompanied by a reduction in expenditure, and hence taxation, on other goods and services. There will be a compensating nonresource element in the remainder of the economy equal to the change in taxation in the nontransport sector. The adjustment consists of estimating the change in expenditure between the sectors ($\Delta T/\phi_t$) and factoring this by an assumed nontransport sector tax rate ϕ_n that, in the absence of any specific knowledge of the alternative consumption foregone, can be assumed to be the average indirect tax rate in the remainder of the economy. Hence

$$\text{Correction} = - \Delta T (\phi_n/\phi_t)$$

where

ΔT = increase in tax paid within the transport sector modeled, and

ϕ_t, ϕ_n = taxation rates in the transport sector modeled and in the remainder of the economy respectively.

Summary of the Estimation Expression

The foregoing sections have indicated that the movement costs and benefits may be estimated as follows:

Total movement benefit = increase in gross value to travelers (as measured by their behavioral costs) (i) and (ii)
- increase in resources consumed in transport (as measured by resource costs) (iii)
- taxation adjustment for rate of tax in remainder of economy (iv)

Algebraically this is

$$\text{Benefit} = \tfrac{1}{2}\underbrace{\Sigma (u_1 - u_2) (q_1 + q_2)}_{\text{(i)}} + \underbrace{\Sigma (q_2 u_2 - q_1 u_1)}_{\text{(ii)}} - \underbrace{\Sigma (q_2 r_2 - q_1 r_1)}_{\text{(iii)}} - \underbrace{\Delta T (\phi_n/\phi_t)}_{\text{(iv)}}$$

where

u_1, u_2 = unit behavioral origin to destination costs for the mode, purpose, etc., under consideration in two alternative situations, incorporating any special or trial valuation of the components;

r_1, r_2 = unit resource costs (also incorporating any alternative valuations), i. e., time costs plus money costs less indirect taxation;

q_1, q_2 = trips;

ΔT = increase in tax paid on transport;

Σ = sum over all modes, purposes, origin-denstination pairs, income groups, etc., under consideration; and

ϕ_t, ϕ_n = taxation rates in the transport sectors considered and in the remainder of the economy.

Note that Σ $(q_2 r_2 - q_1 r_1)$ is the net change in resource, and it may be necessary or more convenient to estimate it in part or whole in an aggregate fashion.

SUMMARY

The objective of this paper is to provide practical guidance to transport planners and analysts by describing procedures in two areas: (a) the evaluation of movement costs and benefits consequent to changes in networks and management policies and (b) the estimation of the generalized behavioral and resource cost functions for links of networks and for origin-destination pairs that are necessary for this evaluation process and for forecasting behavior.

The procedures are designed for use in situations where the change in network or policy is thought to have strong effects on the trip pattern and individual link loadings. This will generally be the case in the consideration of urban schemes and may be the case for major interurban schemes as well; in both situations, there may be considerable changes in the trip matrices, modal split, and routes used.

The emphasis of the paper is on operational methods. There are many theoretical points of potential dispute or of current ignorance; insofar as possible, judgments have been made and procedures determined to cover these.

Individual networks or management policies cannot be sensibly considered in isolation. Comparisons between alternatives are essential for the valid consideration of the individual proposals. One particularly important comparison is between the various possible future systems under consideration and, as a base, the "do-nothing" situation. In this contex, "do nothing" means to include only those changes to the existing situation that are, for all practical purposes, unavoidable between now and the period under consideration. It is essential, in the context of the evaluation analysis described in this paper, that identical land use and socioeconomic assumptions be made for each of the alternatives considered.

One part of the comparative evaluation of alternative transportation networks or management policies requires the estimation of those costs or benefits that arise directly from the changes in costs of movement and the associated changes in the volume and pattern of movement. This paper concentrates on the evaluation of these "movement costs and benefits" and does not consider the estimation of all the other important costs and benefits (e. g., capital and other initial costs, accidents, development consequences, and environment) nor the integration of all these into the decision-making framework.

Individual travelers are often not aware of the true costs of travel by alternative modes and to alternative destinations and, in fact, may not even have an objective assessment of these costs. In this situation behavioral costs, b, are here defined as those costs that when used in appropriate models give the best empirical fits to observed behavior; with this basis there is reasonable satisfaction that such models and costs can be used to forecast patterns of movement in alternative situations. These costs in practice are described as linear functions of the costs (fares, perceived mileage costs, and so forth) and component times for the various stages of the possible journeys.

Because the behavioral costs represent the best available estimate of the individual traveler's "disutility of travel," it is sensible to estimate the benefits, or increase in value, to travelers in terms of these costs. The user benefit is estimated as the extra that travelers would have been prepared to pay over the behavioral costs they experience; this is the concept of "consumers' surplus" and a measure of the net user benefit may be obtained by summing it over all journeys by all users. However, these behavioral costs will often not represent use of resources; for example, they may differ from resource costs, r, because of misperception of outgoings, profits or losses of operators, and taxation (e.g., on fuel). It is necessary then to add to the consumers' surplus received by travelers the amount by which behavioral costs exceed resource costs.

For purposes of evaluating public sector investments or policies, a public authority could examine the effect of alternative values of some items of individuals' behavioral costs—leisure and commuting time, for instance—differing from those of the individuals themselves. This means that the net benefits would have to be adjusted by the use of special behavioral costs for evaluation, u, that contain these chosen values rather than those in the normal behavioral costs, b.

Part of the value received by travelers is transferred through taxation payments to society at large; such payments do not constitute use of resources and therefore are an addition to total social benefit. However, any increase in expenditure on transport would be accompanied by a reduction in expenditure on other goods and services. It is necessary to make an adjustment to allow for the different rates of indirect taxation on transport and other expenditures; in the absence of specific knowledge of the alternative consumption foregone it can be assumed that it attracts the average rate of taxation on final expenditure in the remainder of the economy.

An expression is developed for the estimation of the movement costs and benefits in terms of the b, u, and r costs and the trip matrices that result from the modeling process. An algebraic summary of the benefit expression is given.

The precise way in which the benefit expression and generalized costs are calculated will depend on the level of detail and form of particular studies. Considerable guidance is contained in the paper to aid the transfer from concepts to computation.

REFERENCES

1. Movement in London. Greater London Council, 1969.
2. McIntosh, P. T., and Powell, T. J. London Transportation Study: Some Aspects of the Economic Evaluation Methods. Planning and Transportation Research and Computation Co. Ltd. Symposium, Sept. 1969, pp. 15-19.
3. Tressider, J. O., et al. London Transportation Study: Methods and Techniques. Proc., Institute of Civil Engineers, March 1968, p. 433.
4. The Economic Assessment of Road Improvement Schemes. Road Research Laboratory, Tech. Paper 75, Ministry of Transport, 1968.
5. Wilson, A. G., Hawkins, A. F., Hill, G. J., and Wagon, D. J. Calibration and Testing of the SELNEC Transport Model. Regional Studies, Vol. 3, 1969, pp. 337-350.
6. Harrison, A. J. Some Results From Using Interview Methods in Time Research. Dept. of the Environment, Highway Economics Unit, Time Research Note 6, 1969.
7. Modal Split: Factors Determining the Choice of Transport for the Journey to Work. Local Government Operational Research Unit, Rept. C32, 1968.
8. Quarmby, D. A. Choice of Travel Mode for the Journey to Work: Some Findings. Jour. of Transport Economics and Policy, Vol. 1, No. 3, 1967.
9. Lisco, T. The Value of Commuters' Travel Time: A Study in Urban Transportation. Univ. of Chicago, Dept. of Economics, PhD dissertation, June 1967. (An abridgment of this paper appears in Highway Research Record 245, 1969, p. 16.)
10. Stopher, P. R. Predicting Travel Mode Choice for the Work Journey. Traffic Engineering and Control, Jan. 1968.
11. Planning for the Work Journey: A Generalised Explanation of Modal Choice. Local Government Operational Research Unit, Rept. C67, 1970.
12. Dupuit, J. On the Measurement of the Utility of Public Works. Annales des Ponts et Chaussées, Second Series, Vol. 8, 1844.
13. Hicks, J. R. Value and Capital: An Inquiry Into Some Fundamental Principles of Economic Theory, Second Ed. Clarendon Press, Oxford, 1965.
14. Quarmby, D. A. Transport Planning in a Multi-Resource Context: The Morecambe Bay Barrage. Operational Research Quarterly, Vol. 21, Special Conference Issue, July 1970.
15. Neuburger, H. User Benefit in Transport and Land Use Plans. Jour. of Transport Economics and Policy, Jan. 1971.
16. Research Into the Value of Time. Dept. of the Environment, Time Research Note 16, July 1970. HMSO, London.

DISCUSSION

Herbert Mohring, Departments of Economics, University of Minnesota and York University

I am in broad sympathy with the proposals by McIntosh and Quarmby for improving the arts of predicting and evaluating the effects of alternative transportation improvements. If factors other than travel time could adequately be taken into account, more accurate predictions would likely result. Using consumers' surplus rather than cost change benefit measures would also be highly desirable. I do not say this because I expect that consumers' surplus measures would yield radically different estimates from other sensibly chosen cost change measures. [I very much doubt that the danger with cost-based measures that the authors discuss has ever led to rejection of a highway improvement proposal. The change in user cost with a fixed trip pattern seems almost invariably to be used in such estimation work.] Rather, I espouse the proposal because it would bring transportation benefit estimation closer to the economic analysis used in dealing with formally similar problems, thereby reducing the still regrettably high frequency with which vast nonuser benefits are associated with such phenomena as land value changes and generated traffic.

Unfortunately, though, this paper is not without its shortcomings. The authors have opened several cans of worms without either sorting the contents out very well or, indeed, recognizing how difficult the sorting process would be. It is to three of these unsorted cans of worms that I would like to point.

First, in regard to their discussion of "generalized costs," it certainly is true that trip attributes other than travel time influence travel behavior. It is also true, however, that the relative importance of these attributes varies substantially from individual to individual. Although the authors do mention this point, they fail to come adequately to grips with it.

Consider a consumer who desires to maximize the utility he derives from spending his fixed income on two commodities—a general purpose good, "purchasing power," and trips from here to there. The consumer enjoys trips not for themselves but rather for what happens once he gets there. Indeed, the time he spends in transit is a source of dissatisfaction, not utility. It can be shown that this sort of consumer will allocate his income between trips and purchasing power as if the price of a trip equals whatever fare he pays plus the value he attaches to travel time times the time required per trip. As a number of studies have shown, the value of travel time varies substantially among individuals and is particularly closely related to their income levels.

Suppose, now, that there are two ways of traveling between here and there. One is fast but involves a high fare; the second is slow but is low priced. Except for time and money, the consumer is indifferent between them. If so, he can be expected to choose the mode (say) that involves the lowest price to him. The difference between the money costs of the two modes divided by the differences between their travel times is, in effect, the price he must pay to save a minute's travel time by using the faster mode. If the value he attaches to travel time is more than this price, he will use the fast, high-priced mode; if not, he will use the slow mode. In such a system, reduction in the fare or travel time for one mode will lead some travelers to shift to it from the other mode—note, some travelers, not all.

In brief, accurate prediction of route and mode choice on a transportation system in which routes and modes have different mixes of travel times, money costs, and other attributes requires taking into account the fact that any given route has as many prices (as many generalized costs) attached to it as there are potential travelers. The authors' procedure does not take this fact into account.

Second, a minor point in the paper, but one worth mentioning because of its importance to the literature on modal choice, is that the authors suggest stratifying the population by, inter alia, automobile ownership in predicting travel behavior. Implicit in this suggestion is what seems to me to be an erroneous assumption of causality except in the very short run. A simple but perhaps not unusual example follows: A husband and wife both work. He must use a car on the job. The choice by her to drive rather than to use an available bus is therefore effectively a choice by them to buy a second

car. In deciding on her travel mode, they will presumably weigh the time she will save by driving against the additional money outlay doing so will require. In these calculations, it is worth noting, the relevant money cost is that of owning and operating a second vehicle, not just out-of-pocket operating costs. Here, clearly, automobile ownership is determined by choice of mode, not the other way around. More generally, the choices of mode, automobile ownership, and, indeed, residence and work place location are interdependent. The process is not one in which the last three variables are exogenous in a single equation that serves to determine modal choice.

To do justice to the last group of problems in this paper that I want to discuss would require a complicated and lengthy discussion, far more than I can provide in this limited space. All I can do here is to point out the existence of the problems and assert that solutions to them do exist. [That is, benefit measures closely related to McIntosh and Quarmby's consumers' surplus techniques can be developed that require no more information than do their measures and that avoid the problems to be discussed (17).]

Generally, the amount of a commodity any individual consumes depends not just on its price but also on his income and the prices of complementary and substitute products. If the price of one commodity changes, consumer demands for other commodities and, quite likely, their equilibrium prices also change. Thus, improving one highway and thereby lowering the price of trips on it will likely serve to divert traffic from other highways. The result will be lower congestion and lower trip prices and, hence, benefits to the users of these facilities that are clearly not directly reflected in the demand schedule for the originally improved highway.

An obvious extension to the authors' analysis suggests itself to handle this problem: Add up changes in consumers' surpluses not just for the originally improved highway but also for those that are benefited through traffic diversion. The problem is that the position of the demand schedule for one highway, the area under it, and the measured benefit depend on the price of trips on other highways.

Formally, this extension of McIntosh and Quarmby's proposal involves evaluation of a line integral along some particular path. It can be shown that the value of a line integral will depend on the path chosen to evaluate it unless certain integrability conditions are satisfied. It can also be shown that these conditions are not normally satisfied for the sort of demand schedule dealt with by economists generally and by the authors in particular. This being the case, it is quite possible that, using their techniques, the rank order of two alternative improvements would depend on the specific path used in evaluating their benefit line integrals—clearly an unhappy state of affairs. At least one consumers' surplus type of benefit measurement technique that does not suffer from this disability exists, but it is not that implied by McIntosh and Quarmby's paper.

Reference

17. Mohring, H. Alternative Welfare Gain and Loss Measures. Western Economic Jour., Dec. 1971.

AUTHORS' CLOSURE

We would like to thank Mohring for his remarks. They bear on several theoretical points that, although generally covered in the references supporting the paper, are not fully discussed in the body of what was intended to be a paper for practical guidance.

His points may be split into two main areas: those dealing with the use of generalized cost in modeling and forecasting individual and group behavior and those connected with the calculation of benefit. We accept completely that different people will individually have different generalized costs for the same journey. This is fully recognized in our paper and in the forms of models that are generally used in the United Kingdom for modal split and distribution and for which this paper is intended to help provide inputs. These models do not predict behavior on the basis of minimum cost but rather by the

use of logistic-exponential probability splitting functions, dividing any group at risk among the origins, destinations, and modes available.

Such models give deterministic estimates of aggregate behavior, but they are based conceptually on probabilistic hypotheses about individuals' behavior so that variations in behavior among individuals are implicitly allowed for. In practice, to split the population into subgroups according to, for instance, income groups, levels of car ownership, or journey purposes will move toward homogeneity within groups and will reduce the effects of the overall interpersonal variations. But some variations between individuals within groups will remain and are allowed for.

There are, of course, procedures being developed for joint modeling of mode, automobile ownership, and residence and work place locations that use various concepts of utility, generalized cost, and dynamic behavior, but we did not try to include them in a paper deliberately restricted to generally available and fully practicable procedures. To make firm recommendations on how the automobile ownership-choice of mode relationship should be treated (other than by ignoring it) would be dangerous in view of the relatively limited research that has been carried out in this area.

In suggesting that our procedure should be extended to cover traffic benefited by diversion, we think that Mohring may have missed an important concept of our paper, inasmuch as benefits arising from all changes in travel behavior are counted, as well as benefits from simple cost reductions. We deal with the totality of all trips affected; the integration, or summation, of the benefit expression is over all origin-destination pairs and modes and thus takes account of changes between them on account of cost-induced behavior changes. The generalized cost is the aggregation of total origin-destination costs, not the element of cost on a particular element of highway or transit system; similarly, the origin-destination trip is the "good," not the trip from one end of a highway link to the other. Traffic between other origin-destination pairs benefiting from congestion reduction is thus allowed for.

The issue relating to the path of integration and the form of the benefit expression is very complex, and we accept that it is not fully resolved. The point is briefly discussed in our paper and in some detail elsewhere (15). A summary of the argument is given here. Traditionally the argument in favor of using the trapezium measure as an approximation in the case where only one price falls is as follows:

For those people who continue to consume the same amount of a commodity before and after a price change, the benefit must be exactly equal to their change in expenditure (or, in our case, generalized cost). For those who change their consumption pattern, their benefit per unit cannot be greater than those who do not change their behavior; otherwise, they were in a nonrational position before the price change. Also it cannot be less than zero because, otherwise, they would have moved to a nonrational position after the price falls. If we assume that their benefit lies midway between the possible extremes, we obtain the trapezium measure.

We now come to the case of a multiple price change. We divide expenditure into that which continues to be devoted to the same good and that which is switched from one good to another. The first group, as in the one price case, receives a benefit equal to the change in expenditure. By the same argument as in the sample case we may show that the benefit per item of switched expenditure lies approximately half way between the price changes of the goods between which the switch is made. This in turn can be shown to add up to the sum of the trapzium measures taking the demand diagram for each good separately. Clearly this is an approximation in the same way that the single price change case is an approximation.

Theoretically it has been shown that the problem of the path of integration is closely related to variability in the marginal utility of money and is of the same order of significance. In the context of transportation models normally used, it has been shown that very extreme changes of transport provision and costs are required before the issue becomes of a more than negligible importance. Other measures, which avoid these approximations, require enormous amounts of normally unobtainable data for their computation; we feel that these cannot be considered as available for general and routine use.

NUMERATOR-DENOMINATOR ISSUE IN THE CALCULATION OF BENEFIT-COST RATIOS

Gerald A. Fleischer, Department of Industrial and Systems Engineering, University of Southern California

The application of the benefit-cost ratio method to the evaluation of alternative highway designs and programs is of substantial interest. Several important reference works in this area point out that the magnitude of the ratio will be affected by the category to which a specific consequence is assigned, that is, whether an economic gain will be considered as a benefit (added to the numerator) or as a negative cost (subtracted from the denominator). The writers of these references proceed to justify the specific classification of certain consequences such as roadway maintenance costs and user costs. However, inasmuch as the only relevant issue is whether the ratio exceeds unity, the numerator-versus-denominator issue is without interest. A ratio cannot be altered from greater than unity to less than unity merely by adding (or subtracting) a constant from both numerator and denominator.

• SOME AUTHORS are critical of the benefit-cost ratio method on the grounds that the magnitude of the ratio is dependent on whether a particular economic consequence is considered in the numerator as a benefit or in the denominator as a "negative cost." (Alternatively, one may choose between inclusion in the denominator as a cost and inclusion in the numerator as a "negative benefit.")

This question occurs frequently in problems relating to highway construction and design. In particular, consider three major consequences of new highway construction or improvement: (a) capital costs, (b) benefits accruing to users of the facility, and (c) highway maintenance expenses. The issue raised at this point is whether maintenance expenses should be deducted from road-user benefits numerator) or, conversely, added to capital costs (denominator), inasmuch as each strategy will result in a different benefit-cost ratio (except in the case where the benefit-cost ratio equals unity).

A simple numerical example will serve to illustrate this point. Let X = road-user benefits = 15, Y = facility capital costs = 8, and Z = facility maintenance costs = 5. (All economic consequences are stated in terms of their equivalent uniform annual amounts over the life of the project. Alternatively, they could have been stated in terms of equivalent net present value. Either convention is appropriate to the following discussion.)

Now, in the event that maintenance costs are subtracted first from benefits

$$\text{B:C} = \frac{X - Z}{Y} = \frac{15 - 5}{8} = 1.25$$

If maintenance costs are considered in the denominator,

$$\text{B:C} = \frac{X}{Y + Z} = \frac{15}{8 + 5} = 1.15$$

Sponsored by Task Force on Application of Economic Analysis to Transportation Problems.

The critics claim that the resulting ambiguity makes it difficult, if not impossible, to compare two projects by using the benefit-cost ratio method. For example, suppose that we are considering an alternative (Project II) with a benefit-cost ratio of 1.20. Which, then, is preferable? Project I with B:C = 1.25 (or 1.15) or Project II with B:C = 1.20?

In response to this apparent problem, several writers have attempted to specify precisely where annual expenses ought to be placed. The "Red Book" of the American Association of State Highway Officials, for example, suggests that only road-user costs (and benefits) should appear in the numerator; all other economic consequences of the proposed investment should appear in the denominator (1, p. 14). On the other hand, it has been argued that, "...in terms of economics, economy and cost accounting, it is much more logical to put the repetitive annual cash flows in the numerator and the capital investments in the denominator" (4, p. 149. Although these views are not necessarily incorrect, they are at best much ado about very little; and, at worst, they reflect a serious misunderstanding of the application of the benefit-cost ratio method.

There is only one characteristic of the benefit-cost ratio that is relevant to the decision-making process: Is the ratio greater than unity? Otherwise, the absolute value of the ratio is irrelevant. This comment holds for both positive and negative values of the denominator. That is, the decision rules are:

For denominator > 0, accept if B:C > 1.0; reject otherwise.
For denominator < 0, reject if B:C > 1.0; accept otherwise.

In both instances the critical value of the ratio is 1.0. It is the comparison with this benchmark that leads to the decision.

Returning to our example, let us suppose that the benefit-cost ratio of Project II resulted from the following estimates: X(II) = 24, Y(II) = 20, and Z(II) = 0.

Now, let us determine the preferable alternative, I or II, considering maintenance costs in the numerator (as a negative benefit) or in the denominator. In the first, we note that B:C = 1.25, which, because it is greater than unity, leads us to conclude that Project I is preferable to "doing nothing," i.e., investing elsewhere. But is II preferable to I? To answer this question we note that incremental benefits = 24 - (15 - 5) = 14, incremental costs = 20 - 8 = 12, and ΔB:C = 14/12 = 1.17; thus Project II is preferred. Solving under the assumption that maintenance costs should be included in the denominator, we have incremental benefits = 24 - 15 = 9, incremental costs = 20 - (8 + 5) = 7, and ΔB:C = 9/7 = 1.29; as before, Project II is preferred because the incremental benefit-cost ratio exceeds unity.

To prove that this conclusion holds in all cases, we need only note that our decision rule (for positive denominator) is simply to accept the incremental investment if the resulting incremental benefit-cost ratio exceeds unity; otherwise, reject it. Stated in prior notation, the rules are:

$$\text{If B:C} = \frac{X - Z}{Y} > 1.0, \text{ accept; otherwise, reject.}$$

The alternative formulation is

$$\text{If B:C} = \frac{X}{Y + Z} > 1.0, \text{ accept; otherwise, reject.}$$

These inequalities clearly will lead to identical results; that is,

$$\text{If } \frac{X - Z}{Y} > 1.0, \text{ then } X > Y + Z, \text{ and } \frac{X}{Y + Z} > 1.0$$

This result arises from the fact that the direction of an inequality cannot be reversed merely by subtracting a constant from both sides of the inequality.

Another defense of the "annual expenses in the denominator" convention arises from the assertion that the benefit-cost ratio represents (or ought to represent) a measure of the profitability of the investment. By way of illustration, Winfrey (4) provides the following example with X = 100, Y = 1, and Z = 20. Thus,

$$B{:}C = \frac{X - Z}{Y} = \frac{100 - 20}{1} = 80.0$$

Alternatively,

$$B{:}C = \frac{X}{Y + Z} = \frac{100}{1 + 20} = 4.8$$

"The ratio of 4.8 really has no meaning," Winfrey writes, "because essentially it means that the gross profits were 4.8 times the annual operating expense, and the return on invested capital is not calculated. But the ratio of 80.0 does reflect the size of the net return on the invested capital" (4, p. 150).

The difficulty here is that the benefit-cost ratio is not a measure of return on invested capital. It is not meant to be an index of profitability, merely acceptability. In this illustration, both results lead to the conclusion that the proposed investment is attractive. But nothing more! It does not tell us, for example, that a project with B:C = 80 is preferable to an alternative with B:C = 5.

This principle may be illustrated by another example. Consider these alternatives in competition with another. Designate the two alternatives R and S respectively. Assume the following data for S: X = 120, Y = 20, and Z = 4. Thus,

$$B{:}C = \frac{X - Z}{Y} = \frac{120 - 4}{20} = 5.8$$

Alternatively,

$$B{:}C = \frac{X}{Y + Z} = \frac{120}{20 + 4} = 5.0$$

Summarizing the results of the two sets of calculations gives us the following benefit-cost ratios:

Maintenance Costs	Alternative R	Alternative S
Numerator	80.0	5.8
Denominator	4.8	5.0

The only conclusions that may be drawn from these data are that (a) alternative R is preferable to the base condition because its B:C > 1.0, and (b) alternative S is preferable to the base condition for the same reason. We do not conclude that R is "very good" nor that R is not preferable to S. Indeed, the choice between R and S awaits the following analysis of the differences between alternatives: with maintenance costs in numerator

$$\Delta B{:}\Delta C = \frac{116 - 80}{20 - 1} = 1.9$$

With maintenance costs in denominator

$$\Delta B{:}\Delta C = \frac{120 - 100}{24 - 21} = 6.7$$

The incremental benefit-cost ratios (in both instances) exceed unity, indicating that the incremental investment in S over R is justified. Again, other information about the size of the ratio is irrelevant to the decision problem.

In summary, then, the numerator or denominator problem is an empty issue. As a practical matter, the decision-maker should be concerned only with whether the benefit-cost ratio exceeds unity. The absolute value of the ratio, although it can be affected by the choice of numerator or denominator for certain income-expense items, is not relevant to the choice between alternatives.

REFERENCES

1. Road User Benefit Analysis for Highway Improvements. AASHO, Washington, D.C., 1960.
2. Smith, G. W. Benefit/Cost Ratios: A Word of Caution. Highway Research Record 12, 1963, pp. 77-90.
3. Taylor, G. A. Managerial and Engineering Economy—Economic Decision Making. D. Van Nostrand Co., Princeton, N.J., 1964.
4. Winfrey, R. Economic Analysis for Highways. International Textbook Co., Scranton, Penn., 1969.
5. Wohl, M., and Martin, B. V. Evaluation of Mutually Exclusive Design Projects. HRB Spec. Rept. 92, 1967.

DISCUSSION

Robley Winfrey, Consulting Engineer, Arlington, Virginia

The paper by Fleischer calls attention to two misunderstood or misapplied principles of economic analysis of alternative investment schemes, whether mutually exclusive or independent proposals. One of the principles of the analysis for economy is that it is the difference between proposals that is significant and not the magnitude of the cash flows. The second point often misunderstood is the principle that, when using the rate of return or benefit-cost ratio method, the procedure must be by pairs of alternatives (all possible pairs within the specific alternative projects being considered) by the principle of differences. This principle of comparison of alternatives by their differences is the only procedure that will identify the one alternative that will maximize the total net income, or net benefits.

This paper arrives at the right answer to the numerator-denominator controversy when making a choice from a pair of alternatives or from a series of pairs of alternatives, whether mutually exclusive or independent. If income is to be maximized, the factor to watch is simply that the benefit-cost ratio for a pair of alternatives is 1.0 or more or, for the rate of return method, that the calculated rate of return is at least the minimum attractive rate of return sought. In the book by Winfrey, cited by Fleischer, this principle is adhered to in the table on page 136 but, as stated by Fleischer, forgotten on page 148 in the discussion of the numerator and denominator locations for the annual expense factor.

If we accept the truth that, whether the annual expense factor is in the numerator or denominator, either procedure will indicate the same alternative as having the greater economy, then the magnitude of the benefit-cost ratio is irrelevant, except that it is less than or greater than 1.0.

But there are applications of the B/C ratio method of analysis in which the magnitude of the ratio may be of significance to the decision-maker. Consider the following three applications.

In choosing one alternative out of two or more mutually exclusive alternatives, the decision-maker would like to know whether the benefit-cost ratio between a pair of alternatives was 0.9 or 1.1. These near-to-one ratios for all practical purposes are equivalent to 1.0. The degrees of precision and of certainty in the whole process of economic analysis are rough, so that the final answer should not be used to precise magnitudes. The size of the ratio, when considered against external consequences, may be important to the decision-maker.

A second application where the benefit-cost ratio magnitude may be important is in determining the sensitivity of the discount rate (or other factor) in affecting the ratio. The comparative sensitivity of a 7 percent discount factor and a 10 or 12 percent discount factor cannot be determined by whether the benefit-cost ratio is more or less than 1.0. The analyst or the decision-maker needs to know the relative magnitude of the ratios obtained by using, say, a 7 percent rate and when using a 10 or 12 percent

rate. Table 1 gives data that illustrate that the benefit-cost ratio is sensitive in the numerator position but not in the denominator position. This fact should be known to the analyst and to the decision-maker.

A third application may arise where the magnitude of the ratio is important. Consider application of the benefit-cost ratio method to a single proposal under study. Let X = 100, Y = 5, and Z = 20. The calculated benefit-cost ratios are as follows:

1. X in numerator = (100 - 20)/5 = 16;
2. X in denominator = 100/(5 + 20) = 4.

Table 1. Sensitivity of benefit-cost ratio.

		Present Worth for 20 Years		
Item	Amount ($)	7 Percent	10 Percent	12 Percent
X[a]	100	1,059	851	747
Y[b]	120	120	120	120
Z[c]	30	318	255	224
Calculated benefit-cost ratio				
Z in numerator		6.2	5.0	4.4
Z in denominator		2.4	2.3	2.2

[a]X = annual user benefits. [b]Y = initial capital costs. [c]Z = annual maintenance.

The fact that these two calculated benefit-cost ratios are greater than one is not sufficient information to the decision-maker. It is important for him to know that the answer is either 16.0 or 4.0 before he can judge the effectiveness of the proposed investment. And if that person is concerned about the rate of return on his investment he should be guided by the ratio 16.0, the numerator result.

Some highway engineers may calculate the benefit-cost ratio of a proposed improvement to a highway facility as compared to the existing facility for many independent projects and then use the benefit-cost ratios as an index of priority of construction. This procedure is in theoretical error because each of the independently proposed projects is not compared with each other by the differential procedure.

What are the implications of this practice? What procedure do you recommend for priority selection when there may be several hundred or several thousand projects to compare with each other? At times this priority study is made for just a general guide to what projects or what type of improvements are likely to afford the greater benefits. Actual construction priorities are determined with more refined calculations.

Fleischer is to be commended for giving us this simple and correct explanation to the long-existing numerator-denominator controversy when selecting from a group of mutually exclusive alternatives and in priority selections.

AUTHOR'S CLOSURE

I note with interest Winfrey's three exceptions to the rule that the magnitude of benefit-cost ratio is irrelevant, other than the information that the ratio is less than (or greater than) unity. I concur with the first observation. Clearly, project benefit-cost ratios close to unity suggest that the accept or reject decision may be sensitive to the estimates associated with one or more of the inputs. In this case, sensitivity analyses may be in order.

With regard to the second exception, it is not clear to me that the numerical example supports the conclusion that "the benefit-cost ratio is sensitive in the numerator position but not in the denominator position." In both instances—annual maintenance costs in the numerator or the denominator—the benefit-cost ratios are well in excess of unity. Sensitivity has to do with changes in decisions as the result of changes in input values. The fact that the ratio is reduced from 6.2 to 4.4 as the discount rate increases from 7 to 12 percent is of no particular relevance.

The third exception provided by Winfrey also is not convincing. With regard to the decision, a benefit-cost ratio of 16.0 is equivalent to one of 4.0 inasmuch as both lead to the same result, i.e., accept the proposed investment. Neither ratio indicates the rate of return of this project. In fact, this example is instructive in that it illustrates that a single proposal can result in two entirely different benefit-cost ratios. Yet these ratios are (necessarily) consistent in that they lead to identical decisions.

PASSENGER CAR FUEL CONSUMPTION AS AFFECTED BY ICE AND SNOW

Paul J. Claffey, Consulting Engineer, Potsdam, New York

The effects of road surface ice, hard-packed snow, and various depths of newly fallen snow on the fuel consumption of a typical passenger car were examined during the winter of 1970-1971 on a straight, level test road near Ogdensburg, New York. Data on operation under various conditions of ice and snow were compared directly to data on dry road operation. The results given in this paper include the rate of fuel consumption of the typical passenger car in relation to speed for each of the ice and snow conditions involved in the study, the straight-line relationship between the fuel consumption of the typical passenger car and depth of newly fallen snow for three running speeds, and the factors to correct dry pavement fuel consumption rates for the different ice and snow conditions for the various speeds. The worst ice and snow condition as far as fuel consumption is concerned is snow depth. Fuel consumption will be 50 percent more on a road with a 2-in. snow depth than on a dry road. Information developed in this study and examples of applications in highway economy analyses are presented.

•THE DIFFICULTIES associated with ice and snow are widespread in the United States. Annual snowfalls of 40 in. or more are common in about 40 percent of the total area of the 48 conterminous states. Highways in 22 states, including the populous northeastern states, are subject to the effects of such snow accumulations each year. In many other states there are at least one and perhaps two or three traffic-disruptive snowstorms every winter.

State and local highway departments responsible for maintaining road service in the snow states must decide how frequently roads are to be plowed during and after snowstorms, the tolerable depth of snow on road surfaces for various traffic volumes, and when to use salts rather than sand to remove ice to give traction on surface ice. Because these decisions relate to the operating costs of highway users as well as to the levels of maintenance expenditure that are involved in snow clearing, adequate data on the effect ice and snow have on vehicle operating costs (especially fuel consumption) are needed in the snow states.

EFFECT OF SNOW ON VEHICLE OPERATION

Ice and snow conditions restrict vehicle movement in a variety of ways, depending on the actual condition of the ice or snow on the pavement. Both ice and snow, but particularly ice, cause excess fuel consumption by inducing slippage of the traction wheels, which in turn produces engine revolutions without corresponding vehicle movement. Both ice, when it freezes into shallow ruts, and snow, which packs down into a rough washboard-like surface, present an irregular running surface for vehicles. This wrinkled surface causes vehicles to consume extra fuel because they must continually climb over these irregularities to produce forward movement. Freshly fallen snow of 1 in. or more in depth also increases vehicle fuel consumption because of the effort needed

Sponsored by Task Force on Application of Economic Analysis to Transportation Problems.

to pack down the snow under the wheels as vehicles move along and the necessity to climb over and across ruts left by other vehicles. All ice and snow conditions involve considerable side throw of vehicles at speeds above 30 mph. This also adds to vehicle fuel consumption.

FUEL MEASURING TEST

The effect on fuel consumption of ice and snow on the roadway was determined by direct measurement of the fuel consumed by a typical passenger car operating on a level, straight section of good paved road for a variety of ice or snow surface conditions.

Test Road

The test road was the same section of highway that in previous years had been used for determining fuel consumption for optimum road geometrics in connection with the vehicle operating cost study conducted for the National Cooperative Highway Research Program (1). It is 4,000 ft long and has adequate approach sections and convenient places to turn around. There are no driveway entrances throughout the full distance between turn-around points, and traffic flow is negligible. The road is located in northern New York near Ogdensburg, where over 100 in. of snow fell during the 1970-1971 winter season.

Automobile

The test car was a 1964 Chevrolet sedan with a 283 cu in. V-8 engine and automatic transmission. It weighed 4,000 lb during test operations with all test personnel and equipment aboard. This vehicle was the principal passenger car used in the operating cost study reported by NCHRP (1). Engine performance was satisfactory. The engine consumed fuel under ideal test conditions at the time of the snow study (winter of 1970-1971) at about the same rate that it did during the 1964-1967 period when it was used for obtaining the data for the NCHRP study. Snow tires of a good grade that had previously been used for 5,000 miles of winter travel (typical wear of snow tires) were mounted on the traction wheels for this test program.

Fuelmeter

Fuel consumption was measured with the electronic fuelmeter that had been developed for the NCHRP study. It measures fuel consumption to the nearest 0.001 gal with dependable accuracy under all normal operating conditions.

Procedure

Test operations were carried out over the measured test section both for uniform running speeds of 20, 30, 40, 50, and 60 mph (when practicable) and for stop-and-go speed change cycles at running speeds of 30 and 50 mph. Uniform speed runs were made for each of the following ice or snow conditons:

1. Very slippery hard-packed snow and ice, with about 20 percent of the running track bare road;
2. Hard-packed snow on ice with irregular bumpy surface and wrinkles formed by many vehicles passing over snow left after close plowing;
3. One-half in. of new snow on hard-packed snow (about the depth left after passage of snowplow);
4. Three-quarter in. of new snow on hard-packed snow;
5. One in. of new snow on hard-packed snow;
6. One and one-half in. of new snow on hard-packed snow; and
7. Two in. of new snow on hard-packed snow.

Data were obtained for operations under conditions 3, 4, 5, 6, and 7 during the same snowstorm so that all factors affecting results were identical except the snow depth.

Stop-and-go speed change cycles were made only on very slippery hard-packed snow and ice (condition 1). All test runs were made at air temperatures between 25 and 30 F and only during periods of calm wind conditions.

Fuel consumption was also measured for test runs on dry pavement during late fall of 1970 and early spring of 1971. This was done for the following reasons: (a) to provide comparison data for analysis of the increase in fuel consumption due to ice and snow conditions and (b) to serve as a check on the accuracy of the fuelmeter and the constancy of the basic fuel consumption characteristics of the test car's engine. Because the air temperature during the dry road test runs (40 F) was necessarily somewhat higher than that during the ice and snow test runs (25 to 30 F), a correction factor from NCHRP Rept. 111 (1, p. 63) was applied to adjust dry road fuel consumption data for the lower temperature.

Results

The relationship between fuel consumption rates and vehicle speeds for passenger car operation on roads with the various ice and snow conditions is shown in Figure 1. Curve A shows the fuel consumption for the test vehicle operating on dry pavement. Data for this curve were obtained by operation at ambient temperatures higher than those encountered during the ice and snow test operations (40 F compared to 25 to 30 F). However, before plotting curve A, these data were corrected to what they would have been at 25 to 30 F by using temperature curves mentioned earlier (1, p. 65). Test operations were carried out at speeds up to 60 mph on bare pavement, on ice-covered roads, and for snow depths up to ½ in. However, for snow depths greater than ½ in., maximum test speeds were limited to 50 mph largely because of the severe side throw drivers encountered in deeper snow when traveling at high speeds.

The excess fuel consumed for stop-and-go cycles on very slippery, hard-packed snow and ice (condition 1) was found to be 0.008 and 0.017 gal per stop for 30- and 50-mph running speeds respectively. These values are close to those observed for stop-and-go cycles on dry pavement, 0.010 and 0.017 gal per stop. However, the excess time consumed for stop cycles at 30 and 50 mph is approximately 50 percent greater on ice- and snow-covered pavement than on dry pavement. Apparently any extra fuel consumption due to slipping on the ice during the acceleration portion of the stop cycle is compensated for by reduced consumption due to lower acceleration on ice.

The curves of Figure 1 show that the ice and snow condition having the most severe effect on passenger car fuel consumption is newly fallen snow. Even as little as ½ in. of snow (curve 3) will induce fuel consumption rates greater than either a very slippery, hard-packed snow surface (curve 1) or a less slippery, but bumpy, wrinkled surface (curve 2). Curves 3, 4, 5, 6, and 7 give the fuel consumption rates for road conditions that are identical except for snow depths, which are ½, ¾, 1, 1½, and 2 in. respectively. It is evident from Figure 1, however, that all roads with ice or snow or both, whether principally slippery, rough, or snow-covered, produce a substantial increase in passenger car fuel consumption compared to operation on dry pavement.

The curves of snow depth versus passenger car fuel consumption for 30-, 40-, and 50-mph running speeds are shown in Figure 2. These curves show that the effect of snow depth on fuel consumption increases with increases in speed. The principal reasons for this increase are the side throw and rough handling experienced by drivers traveling at high speed over the ruts left in a fresh snowfall by other vehicles.

Table 1 gives correction factors to adjust passenger car fuel consumption rates on dry pavement for operation when the road surface is covered with ice and snow. If the dry surface fuel consumption rate of a particular type of automobile (or of passenger cars in general) is known for travel on a road having given geometrics, the fuel consumption when the road is covered with any of a variety of snow or ice conditions may be found by multiplying by the appropriate correction factor from Table 1. Dry pavement fuel consumption rates should be corrected for temperature before applying correction factors for snow conditions.

Figure 1. Fuel consumption rates of a passenger car for various ice and snow conditions.

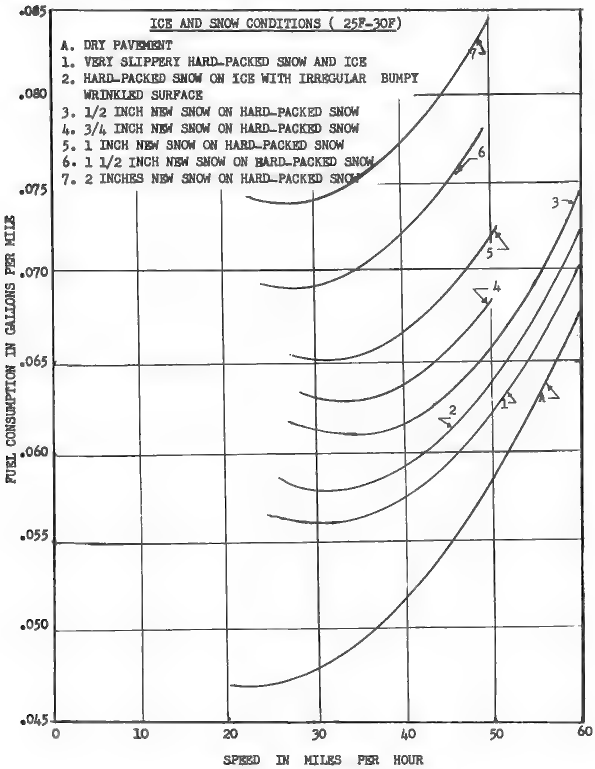


Figure 2. Relationship between snow depth and rate of fuel consumption of passenger cars.

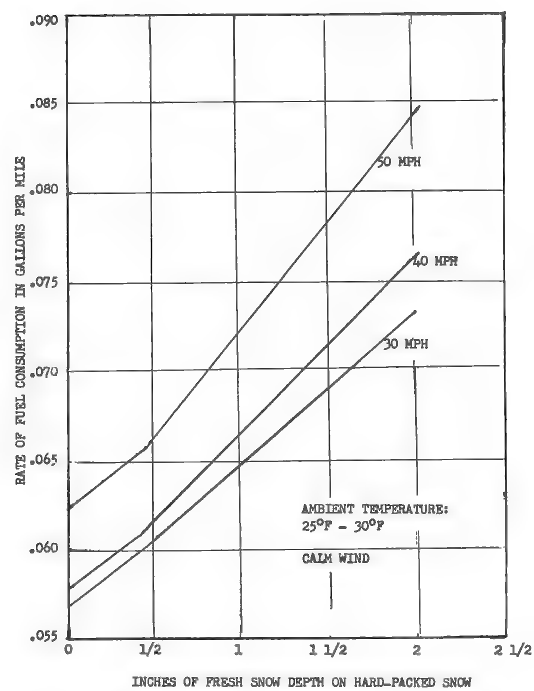


Table 1. Correction factors to adjust passenger car fuel consumption for ice and snow conditions.

Speed (mph)	Dry Pavement	Very Slippery Hard-Packed Snow	Hard-Packed Snow on Ice With Bumpy Surface	New Snow on Hard-Packed Snow (in.)				
				½	¾	1	1½	2
20	1.00	1.23	1.30	1.36	1.43	1.47	1.51	1.60
30	1.00	1.16	1.20	1.28	1.32	1.35	1.45	1.54
40	1.00	1.11	1.14	1.20	1.23	1.28	1.40	1.48
50	1.00	1.06	1.10	1.12	1.18	1.24	1.34	1.45
60	1.00	1.04	1.08	1.10	–	–	–	–

Note: Correction factors are designed to be applied to values in Table 6 of NCHRP Report 111. They may, however, also be applied to any valid passenger car fuel consumption rates for operation on dry pavement.

APPLICATION OF RESULTS

The results of this research have many applications in highway engineering economy analysis, including the following:

1. Evaluation of the extra cost to road users of operating on ice or snow in order to justify the cost of accelerated ice and snow removal;
2. Comparison of total passenger car fuel consumption costs over alternate routes where one is subject to substantial ice and snow cover and the other is free of snow problems;
3. Determination of spacing of gasoline service plazas along limited-access roads subject to ice and snow;
4. Prediction of fleet fuel consumption costs when operations are in regions where roads are snow-covered for part of the year; and
5. Selection of geometric design details for roads in snow areas to compensate road users for the extra operating costs incurred because of snow conditions.

SAMPLE PROBLEM 1—JUSTIFICATION OF SNOW REMOVAL PROGRAM

Problem

A nearly level, two-lane, two-way, high type of asphalt road interconnects two small cities 6 miles apart. During the normal workday rush period (4:00 to 7:00 p.m.) an average of 3,000 passenger cars move from one city to the other over this road. On a particular winter day a heavy snowfall occurring between 3:00 and 8:00 p.m. is able to maintain a 2-in. depth of new snow on a layer of traffic-packed snow. Assuming that gasoline costs $0.40/gal and that vehicles maintain an average speed of 50 mph, determine the total fuel cost of the 3,000 cars using the road during the snowstorm (a) if no attempt is made to remove the snow, (b) if plowing is continuous during the peak hours, thus maintaining a hard-packed snow surface, and (c) if only limited plowing is provided so that 50 percent of the cars encounter a hard-packed snow surface while the other 50 percent encounter an average of only 1 in. of new snow on a hard-packed snow base.

Solution

Compute the fuel consumption of 3,000 passenger cars operating at 50 mph on a level high type of pavement for a distance of 6 miles for each of the given ice and snow conditions [the fuel consumption rate on level, dry pavement is 0.052 gallon per mile (gpm)].

If the surface is covered with a 2-in. layer of fresh snow lying on hard-packed snow (no plowing), the correction factor for this condition (Table 1) is 1.45. The fuel consumption rate for this snow condition (0.052 gpm × 1.45) is 0.075 gpm. Total fuel consumption (0.075 gpm × 3,000 cars × 6 miles) is 1,350 gal. The total fuel consumption cost (1,350 gal × $0.40/gal) is $540.00.

If the surface is maintained as hard-packed snow by complete plowing, then the correction factor for this condition is 1.10. The fuel consumption rate (0.052 gpm × 1.10) is 0.057 gpm. Total fuel consumption (0.057 gpm × 3,000 cars × 6 miles) is 1,026 gal. The total fuel consumption cost (1,026 gal × $0.40/gal) is $410.40.

If 50 percent of traffic travels on a hard-packed snow surface and 50 percent on a 1-in. thickness of fresh snow (limited plowing), two correction factors are used: hard-packed snow surface, 1.10; 1 in. of snow on hard-packed snow, 1.24. The fuel consumption rates are 0.057 gpm (0.052 gpm × 1.10) for hard-packed snow surface and 0.064 gpm (0.052 gpm × 1.24) for 1 in. of snow on hard-packed snow. Total fuel consumption is 513 gal (0.057 gpm × 1,500 cars × 6 miles) for hard-packed snow surface plus 576 gal (0.064 gpm × 1,500 cars × 6 miles) for 1 in. of snow on hard-packed snow, which equals 1,089 gal. The total fuel consumption cost (1,089 gal × $0.40/gal) is $435.60.

Summary

1. User fuel cost with no plowing is $540.00.
2. User fuel cost with full plowing is $410.40.
3. User fuel cost with limited plowing is $435.60.

The user fuel saving with full plowing ($540.00 - $410.40) is $129.60. User fuel saving with limited plowing ($540.00 - $435.60) is $104.40. The cost to have a snowplow drive back and forth between the two cities continually during the rush hours is amply justified by the fuel cost savings experienced by the users.

SAMPLE PROBLEM 2—USER COST COMPARISON FOR ALTERNATE ROUTES

Problem

Two through routes connect cities A and B. Each is a level, two-lane, two-way, high type of asphalt road. One route is 500 miles long but runs along a lakeshore where heavy snowfall is frequent during the winter. The alternate route is 600 miles long but by swinging southward escapes serious snowfall. A driver contemplating a trip between these cities is advised that the lakeshore route is covered by 2 in. of fresh snow for the entire length and will not be plowed before his trip is made. The longer route, however, is free of all snow. The trip on the lakeshore route would be made at an average speed of 50 mph whereas speed on the alternate route would be 60 mph (trip time is 10 hours in each case). On which route would fuel consumption cost be least (assuming a fuel cost of $0.40/gal)?

Solution

Compute the fuel consumption cost to operate the passenger car on each of the alternate routes.

1. On the lakeshore route: The fuel consumption on dry pavement at 50 mph (1, p. 17) is 0.052 gpm. The correction factor for 2 in. of snow is 1.45. The fuel consumption rate on the lakeshore route (0.052 × 1.45) is 0.075 gpm. Fuel consumption for the entire trip (0.075 gpm × 500 miles) is 37.5 gal. Fuel cost for the trip (37.5 gal × $0.40/gal) is $15.00.
2. On the southern route: Fuel consumption on dry pavement at 60 mph is 0.058 gpm. Fuel consumption for the entire trip (0.058 gpm × 600 miles) is 34.8 gal. Fuel cost for the trip (34.8 gal × $0.40/gal) is $13.92.

Summary

The southern route is the least costly route for the trip. Fuel cost on the lakeshore route would exceed that on the southern route by $1.08 ($15.00 - $13.92).

CONCLUSION

Ice and snow conditions have a direct impact on passenger car fuel consumption. Slipperiness, as such, does not add substantially to fuel consumption because drivers tend to use extra care to control speed on ice and to hold wheel slippage to a minimum. However, road ice is usually frozen into an irregular surface as a result of the frequent passage of vehicles during the freezing process. This roughness induces substantial extra fuel consumption.

The most severe increase in passenger car fuel consumption due to ice and snow conditions arises when new-fallen snow depths of more than ½ in. are allowed to accumulate. In this situation vehicles need extra fuel not only to pack down snow beneath the wheels but also to propel vehicles across and over surface irregularities due to rutting by previous vehicles. A 2-in. layer of snow on a section of road may add 50 percent to the dry surface fuel consumption of passenger cars.

Ice and snow conditions are unavoidable during the winter season on many highways. Road users and those responsible for maintenance on these roads should be aware of the effects, often severe, that ice and snow can have on vehicle fuel consumption.

REFERENCE

1. Claffey, P. J. Running Costs of Motor Vehicles as Affected by Road Design and Traffic. NCHRP Rept. 111, 1971, 97 pp.

LIFE CYCLE COST AS A CRITERION FOR OPTIMIZING THE CAPACITY OF VEHICLE TERMINALS

Jason C. Yu and Wilbert E. Wilhelm, Jr., Department of Civil Engineering, Virginia Polytechnic Institute and State University

The goal in planning and developing transportation terminal facilities is to provide capacity adequate to meet most demand. Capacity should be such that no substantial portion goes unused for so much of the time that the facility becomes uneconomical. Thus, some caution must be exercised in determining the optimum capacity for the normal period of demand pattern. This study was an attempt to develop a model that would specify the optimum vehicle storage capacity of a typical terminal facility over its entire life cycle. In approaching this objective, simulation methodologies, economic analysis, and statistical methods were blended and directed toward a practical solution of the problem. Specifically, the economic trade-off and the level-of-service concept were used to assist in the cost-effectiveness analysis. For illustration purposes, a microscopic model describing an individual terminal of the minicar transit system was formulated, tested, and refined. The resulting model is intended to be general and flexible enough to be used in planning the terminal capacity of any transportation mode.

•WITH the steady increase in traffic volume on highly congested urban streets, many potential solutions to the resulting transportation problems have been examined, including alteration of the existing travel mode and completely new transportation systems. Many of these innovations have been presented in experimental form in an attempt to reduce urban street congestion. Yet, the urban traveler still prefers the comfort, convenience, flexibility in routing, and manageable cost of his own private vehicle. Transportation system studies (1, 8) have examined the feasibility and desirability of introducing a system of small, electrically powered vehicles (minicars) into the highly populated urban area. This system provides users with the direct benefit of the standard private automobile and, at the same time, reduces urban congestion, noise, and pollution. A fleet of these small vehicles collects and distributes people on a rental basis. The proposed operating system would restrict the minicar movement between specially designed terminals. A user would rent a vehicle at the terminal nearest his origin, drive to the terminal nearest his destination, and leave the vehicle at the destination terminal. A large number of terminals would be provided either through adaptation of existing parking facilities or by construction of new ones. The terminal would be used both for vehicle storage and as a system access point.

A study by Yu (13) examined the improvement in parking space utilization when the minicar system is introduced. It was concluded that the ability of this system to pack more cars into the given amount of parking space can strongly and favorably influence the urban parking situation. Another interesting aspect of this system would be the determination of the optimal storage capacity of terminals. If each point served by a terminal reaches a peak accumulation of parked vehicles at the same time each day, the terminal would have to be sized at a uniform maximum. However, this is not

---

Sponsored by Task Force on Application of Economic Analysis to Transportation Problems.

realistic because all points will not reach a peak simultaneously so that either the timing functions of all demand points within an area served must be known or else the demand of all points at the time the terminal sees its peak load must be known. Also, the duration of vehicle storage for each individual trip also has significant bearing on the required terminal capacity.

The minicar transit mode would have many characteristics similar to the automobile, inasmuch as only licensed drivers use both modes. Therefore, arrival rates and departure rates for the two modes will have comparable characteristics. However, several important factors will differ. A previous study (14) showed that each minicar terminal would provide service to a bounded area within the CBD. Thus, the terminal must be expected to serve a particular set of customers whose trips either originate or terminate within this area near the terminal. Because all minicars will be alike, customers will be indifferent as to which one they use. A customer may select any available minicar when departing from the terminal. A particular minicar, therefore, will not be parked in the terminal until a certain customer returns to find the car for his departure trip. Most automobile parking facilities charge a graduated fee based on the parking duration. Because of the indifference between minicars, a fixed, prorated portion of the total rental cost of the minicar may be allocated to the operation of terminals. Because the fleet system will require a large initial investment and because it is intended as a benefit to society in general, it has been proposed that such a system be financed with public funds (1, 8). If this were done, it would be more appropriate to optimize the system on the basis of minimum total cost rather than maximum profit.

The main objective of this study was to provide a solution technique to determine the optimum capacity of minicar terminals. Because this optimum is found by balancing the cost of waiting to enter the terminal when it is filled to capacity with the cost of providing additional parking spaces in the facility, the procedure makes allowance for some planned waiting. Further objectives of this study were to determine the optimum capacity of a facility over its entire life cycle by considering changes in both the demand and total investment cost for a life cycle terminal capacity. Although the solution method developed was directly applied to the minicar transit system, the basic model with only minor changes should be applicable to other systems, such as automobile parking facilities, seating capacity on buses and trains, and other service facilities.

METHOD OF APPROACH

The study was concerned with determining the optimum economic capacity. There are occasional periods of high demand for which it would be uneconomical to provide quick and easy accommodation to every customer. To accommodate this demand without losing customers' goodwill would require a vast amount of parking spaces, most of which would be unused during the rest of the time period. In other words, when capacity of a facility is planned, it is assumed that occasional and infrequent periods of some overload will occur and will be tolerated. One solution to this dilemma may be the inclusion in the design of the ability to expand without rebuilding the entire facility to accommodate growth if and when needed.

Facility operation may be viewed as a queuing system in which the server (facility) is capable of serving a number of parked vehicles at any time and the service rate is the vehicle parking duration. Arrivals and services are stochastic processes that, for each trip purpose, are functions of time of day and day of the year. If the facility is filled to capacity at the time of arrival of a vehicle, the vehicle must enter a queue and wait for a storage space to become available. For an individual facility, the queue may build along the aisles within the facility or at the entrances to the facility.

Demand Representation

Demand patterns for the minicar terminal are difficult to forecast because the vehicle represents an innovation to transportation, and no historical trends on which to base predictions exist. However, the mode has many characteristics similar to the standard private automobile, and these were used to model the demand for the terminal under study.

Trip purpose, periodical demand fluctuations within a year, and year of the terminal life cycle have all had an important influence on the arrival rate (demand) of vehicles at the terminal. This study attempted to represent these fluctuations in demand as they might actually occur. The arrival process was assumed to be a Poisson distribution because it appears to fulfill all basic assumptions on which the Poisson process is defined (5):

1. The process $\{N(t),\ t > 0\}$ has independent increments.
2. In any small interval there is a positive probability that an arrival will occur, but it is not certain that an arrival will occur.
3. In sufficiently small intervals, at most one arrival can occur.
4. The process has stationary increments within each hour of day. Daily arrivals actually are distributed according to a nonhomogeneous Poisson process inasmuch as the mean rate of arrivals changes each hour of day.

In this study, four trip purposes that would correspond to local trips in a minicar system were assumed: work, shop, business (each to and from the CBD), and the intra-CBD trips (including all trip purposes). Each trip purpose was assigned a different arrival rate (mean of a Poisson distribution) over each hour of the day so that 40 distributions were used over each 10-hour day by the four trip purposes. The seasonal fluctuations were modeled by applying a daily adjustment factor to each distribution to adjust the mean of all arrival distributions within each day. Although this factor is actually a function of the day within a year, the study used a normal distribution and selected the adjustment factor as a random deviation from this distribution for each day. Year-to-year fluctuation was represented by a similar factor based on the assumed growth rate.

The duration of time that each minicar spends in the terminal depends on the departure demand for vehicles at a particular facility. To reflect this random process required that the parking duration for a given arrival be a random deviation generated from a normal distribution, where different normal distributions were assumed for each trip purpose and hour of day according to the arrival time. Because all minicars are alike, a customer uses any available minicar when leaving the terminal. For any given minicar, then, the arrival process is independent of the departure process. Independent departure rates could be used instead of the approach used in this study.

Figure 1 shows the assumed mean arrival rates by hour of day for each trip purpose. The arrival rate for each trip purpose reaches a peak each day; however, the peak time varies among trip purposes. Figure 2 shows the normal distributions that were used for vehicle parking duration for those vehicles that arrived the first hour of each day. These distributions were truncated to provide realistic parking times (allowable range was from 2 min to 12 hours).

Definitions of Costs

As indicated previously, terminal capacity was optimized with the objective of minimizing total costs. By definition, the total cost is the sum of the cost of providing and maintaining the terminal, plus the cost of waiting for a parking space when the terminal is filled to capacity at the time of an arrival. Mathematically, the objective function may be represented as

$$TC = (FC) + (VC)\,(CAP) + \sum_{i} (VT)_i(WT)_i$$

where

TC = total cost,
FC = fixed terminal cost,
VC = variable terminal cost,
CAP = terminal capacity,
VT = value of waiting time by trip purpose,
WT = total waiting time by trip purpose, and
i = trip purpose.

Figure 1. Mean arrivals by trip purpose and by hour of day.

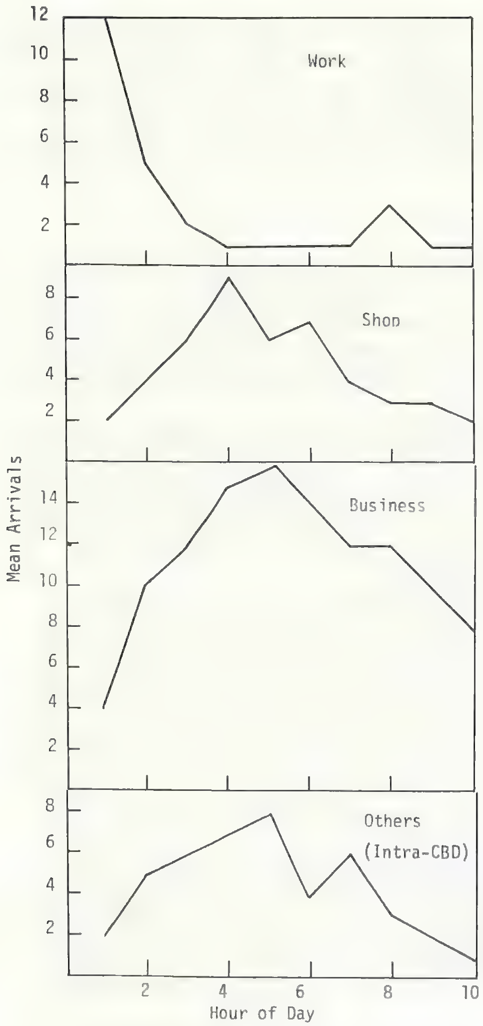


Figure 2. Vehicle parking duration (in minutes) for first hour of day.

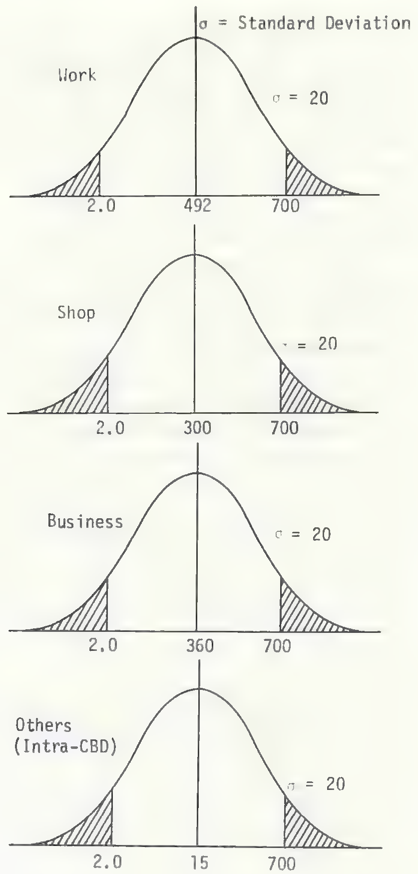
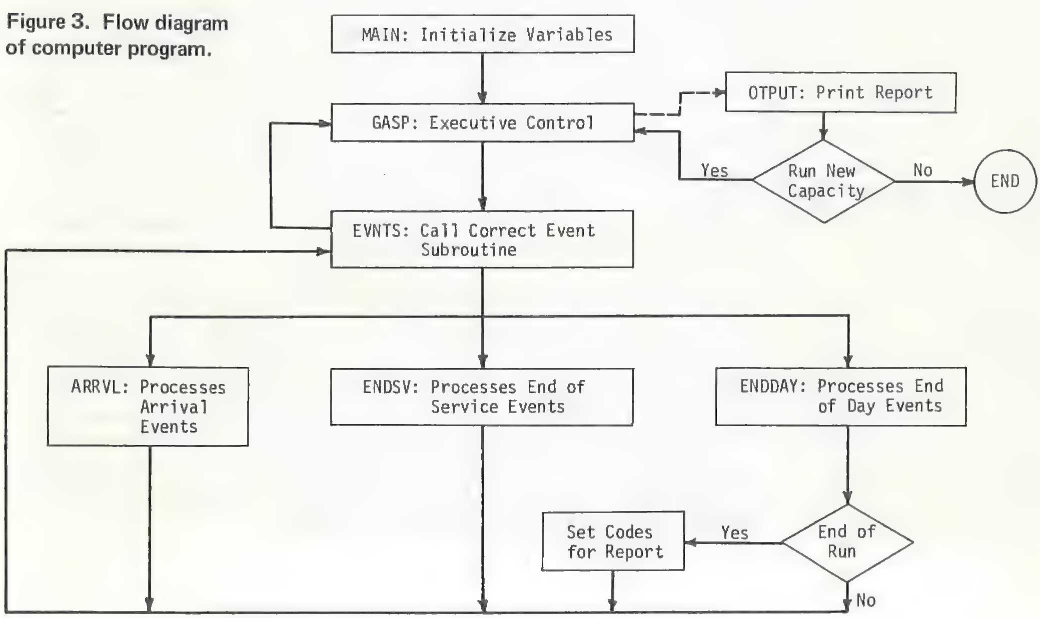


Figure 3. Flow diagram of computer program.



This study employed a fixed terminal cost of $20 per day and a variable cost of $0.46 per space per day, based on a previous study (14). The value of waiting time is a function of trip purpose because it is logical to assume, for instance, that a businessman will place a higher value on his time than a shopping housewife. The values assumed were $2.50, $2.00, $3.50, and $2.75 per hour for work, shop, business, and other trips respectively.

Simulation Program Description

Simulation techniques were applied to study the characteristics of the parking facility inasmuch as many important facets of the system may easily be described through simulation. Analytical procedures were not used to define the system because of the difficulties in problem formulation. The system never reaches a steady state, and the daily arrival rate is a nonhomogeneous Poisson process (nonstationary increments) due to the difference in hourly arrival rates.

It was felt that the discrete-event philosophy of simulation is most suitable to obtain the definition of the system under study. This simulation concept maintains that a system remains static until an event occurs that may cause a change in the state of the system. When an event occurs in the simulated time, only the effects of that particular type of event need be modeled. The GASP simulation language (6) was used because it provides an efficient means of discrete-event simulation.

GASP is essentially a set of FORTRAN-coded subroutines that provide necessary functions for simulations: executive control, gathering of statistics, generation of random numbers from a variety of probability distributions, dynamic storage of variables, and generation of reports. GASP maintains a file of events that will occur and will cause the appropriate subroutine to process an event when it occurs in simulated time. Only three types of events are necessary to model the terminal system: an arrival of a minicar at the facility, a departure from the facility, and an end-of-day event. At the occurrence time of an event, GASP removes the event attributes (or characteristics) from the event file, sets the code for the type of event (one of the attributes), and calls the appropriate subroutine to process the event.

As shown in Figure 3, the simulation program for this study consisted of a main program (that merely initializes values of variables and calls the GASP package), the EVNTS subroutine, the three event processing subroutines (ARRVL, ENDSV, and ENDDAY), and the OTPUT subroutine that provides a special report of the economic analysis for this problem. EVNTS merely calls the correct programmer-supplied subroutine to process that type of event.

Subroutine ARRVL processes all possible changes to the state of the system by an arrival at the terminal. The hour of day of the arrival is calculated, and the total demand variable is incremented by one. The next arrival event is then generated and stored in the event file to occur at some later time. Because the number of arrivals is Poisson-distributed, the time between arrivals is exponentially distributed (5) with a mean equal to the reciprocal of the associated Poisson mean. The appropriate Poisson mean is determined by the hour of the arrival; this value, along with the seasonal and yearly adjustment factors, is used to generate an exponential random deviation that represents the interarrival time. This value is added to the current arrival time and represents the time at which the next arrival will occur. This event time, along with the arrival code, is stored in the event file. The current arrival is then processed. The trip purpose is randomly generated based on the mix assigned for that hour of day; the parking duration for this vehicle is generated as a random deviation from the appropriate normal distribution as described previously. A check is then made to determine if the facility is currently filled to capacity. If a parking space is available, the number of parked vehicles is incremented by one (adding the vehicle to the lot), and an end-of-service event characterized by the time of occurrence (current time plus duration) and the end-of-service code is stored in the event file. Then ARRVL returns to GASP, which causes the next event to be processed. If the facility is full at the time of the current arrival, the attributes of this arrival (arrival time, trip purpose, and duration) are stored in a queue file, and ARRVL returns to GASP.

Subroutine ENDSV is entered to process the removal of a vehicle from the terminal. The total number of parked vehicles is decremented by one, and the number of departures for the current hour of day is incremented. A check is then made to determine if a vehicle is in the queue. If no vehicles are waiting to enter the facility, ENDSV returns to GASP. If there is a queue, the vehicle that arrived first is removed from the queue file, and an end-of-service event is created and stored in the event file. Statistics are then gathered on the time this vehicle waited in queue, and ENDSV returns to GASP for further processing.

Subroutine ENDDAY is entered at the end of each day (every 10 hours). All vehicles in the facility at that time are departed, and queues are emptied if there is a queue at that time. Various daily statistics are then collected by using appropriate GASP subroutines. The first arrival event for the next day and the next end-of-day event are generated and stored in the event file. Other system variables are initialized to start the next day. A check is then made; and, if the simulation run has not completed 250 days (weekdays per year only), ENDDAY returns to GASP to process the next day. If 250 days have been simulated, several codes are reset to cause GASP to print reports.

After GASP prints standard reports, subroutine OTPUT is called to generate a report on the economic evaluation for the terminal. Terminal capacity by year of the life cycle is then incremented, and a sequential run is initiated to simulate the revised terminal capacity. This procedure is used to evaluate the total cost function over a wide range of terminal capacities and years within the life cycle.

SINGLE-YEAR CAPACITY ANALYSIS

A previous study (15) indicated that each terminal should serve a diamond-shaped service area to minimize average walking distance along the rectilinear walking paths found in most cities. Because the minicar transit system must consist of a network of terminals to provide service to urban travelers, the capacity of each terminal depends on the vehicle storage demand generated within the bounded service area. The level of (attracted) demand used in this study represents the part of the total available demand in the service area that is attracted to the minicar transit system. As indicated previously, four trip purposes were assumed in this study. For the intra-CBD trips, total arrivals may be different from total departures, with each customer using only one-way service to or from a terminal. For simplicity, all purposes were handled the same by assuming that the arrival of a vehicle triggered the departure of some vehicle at some future time. It could be assumed that these two processes are independent. The advantage gained in using this approach is that input to the facility equals output.

SINGLE-YEAR SIMULATION

This section describes the important results of simulating demand for a period of 1 year. Various characteristics of the system were defined by the results and the optimum terminal capacity was determined. For each capacity tested, the simulation period covered 250, 10-hour weekdays or the equivalent of 1 year of operation.

The total number of arrivals per hour appears fairly uniform because the input data were arbitrarily selected. The trip purpose mix within the total is, however, quite different from hour to hour as seen in Figure 1. The total number of departures per hour is high during the last half of the day, reaching a peak during the evening rush hour, as expected. This represents a large number of vehicles entering the traffic stream during peak congestion time, and the capacity of the bordering streets should be checked to ensure that it is adequate to handle this increased traffic volume. Figure 4 shows these relationships.

The average number of parked vehicles at the end of each day was found to be about nine. As expected, no units were in the queue at the end of each day. In an actual case, some may park overnight; or, in the minicar system, redistribution at night may cause some vehicles to be in the facility at the beginning of the next day. These vehicles, however, would not affect the waiting times incurred if they depart the facility early the next day before the facility is filled to capacity.

Figure 4. Mean arrivals and departures by hour of day at optimum capacity.

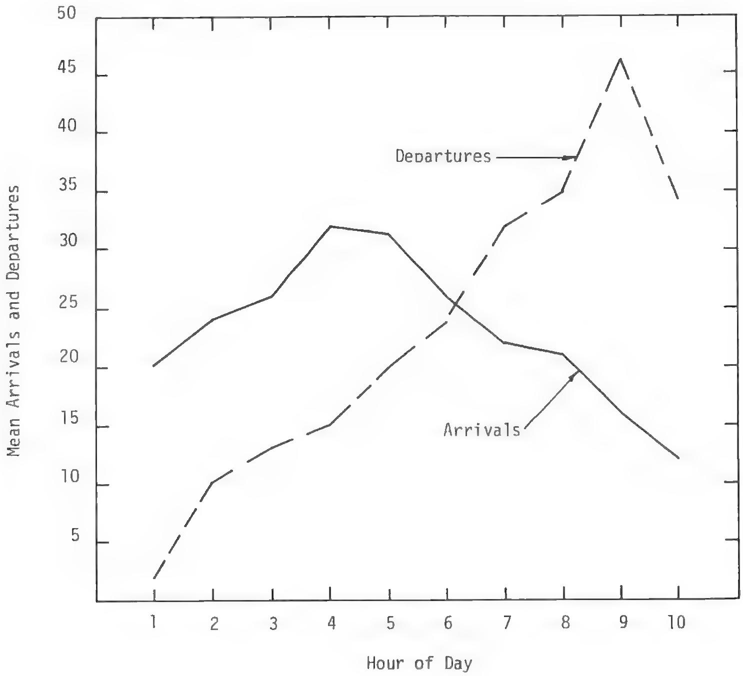
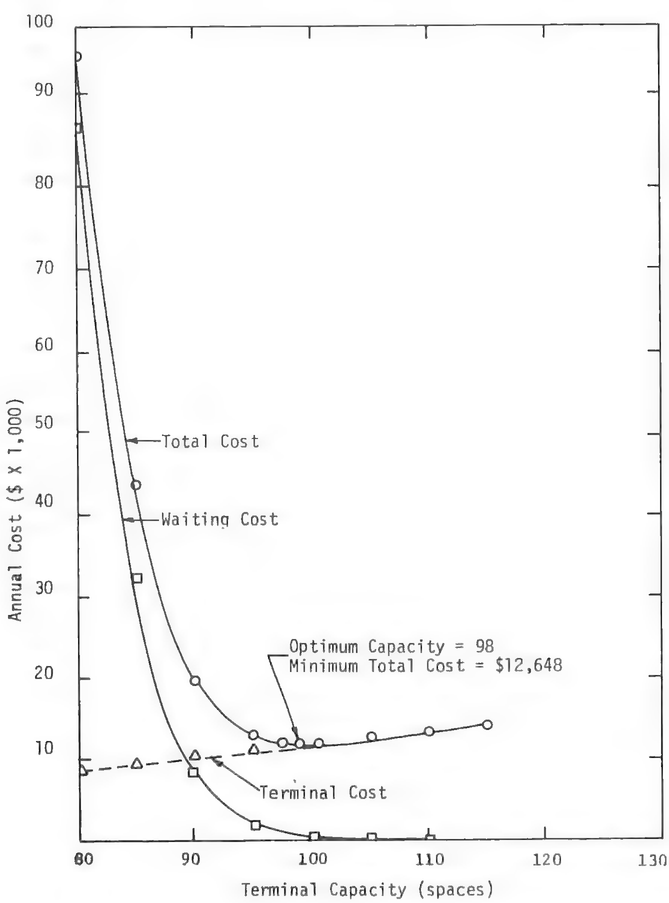


Figure 5. Annual cost versus capacity (1-year analysis).



Peak accumulation of parked vehicles is, of course, a function of the capacity and increases as capacity increases. The average peak accumulation is somewhat lower than the capacity because the peak accumulation will not fill the facility on some days. The expected time of day that peak accumulation was experienced was approximately 1:00 to 2:00 p.m., which corresponds well with results from studies of automobile parking facilities.

Figure 5 shows the expected costs (on an annual basis) as a function of facility capacity. It is seen that small changes in capacity (below optimum) have a large effect on the waiting cost. For the demand rates assumed, minimum total cost occurred at a capacity of 98 spaces. Beyond this capacity, total costs again increase due to increasing facility costs, whereas waiting costs (practically) reduce to zero. For very large capacities, total cost equals facility cost and increases linearly. Table 1 gives the values of the various costs for several capacities. Total cost per year at optimal capacity is $12,648, waiting cost is $298, and facility cost is $12,350.

The total daily demand is described by a mean of 236 vehicles and a standard deviation of 28 (minimum = 168, maximum = 317). This is perhaps low for an actual facility, but demand rates were selected to lower required computer storage space for this study while indicating important system characteristics. The turnover rate averaged 2.41 per day with a minimum of 1.71 and a maximum of 3.23.

At this optimum capacity, 533 vehicles out of a total of 59,004 arrivals waited. The probability that an arrival will wait was therefore 0.009. The expected waiting time of those that waited was 12.5 min. Waiting occurred only on 19 days out of 250 (7.6 percent). The total waiting time was 110.6 hours with the following breakdown: workers, 6.1 hours; shoppers, 23.3 hours; business, 17.2 hours; and others, 64 hours.

Waiting time followed a negative exponential form with only a few vehicles waiting longer times. Figure 6 shows the resulting relative frequency distribution of waiting time. As capacity was increased, the maximum number of units in the queue decreased but at a decreasing rate. This appears reasonable inasmuch as it was shown that, if a vehicle waits in a queue, it will probably wait only a short time. Each increment in capacity, therefore, included a "smaller increment" of queue. Although it may be economical with respect to the parking facility total cost to incur some waiting, there are additional practical factors that must be considered. First, not everyone will wait for a space to become available; balking and reneging may occur even though a customer is supposed to use a given terminal. Consideration of this effect would be included in an expanded study that would include the impact of these changes on neighboring terminals. Second, there must be adequate physical space for the queue of vehicles to build. This might be in the traffic lanes within the facility or along adjacent streets. Under the assumptions of this study, the maximum number of vehicles in queue at one time was 27, the average number was 0.0441, and the standard deviation was 0.7393. The length of the queue is thus seen to be a problem on only a few days during the year. A capacity of 110 was required to eliminate all waiting; this would require investment in 12 additional spaces in the facility.

To study the profit potential of the facility, we assumed that a portion of the minicar rental cost would be allocated to the facility. A value of $0.35 was used for all vehicles, inasmuch as duration will not affect this fee structure for the minicar operated on a fleet parking basis. The profit analysis indicated that profits are lowered as the capacity is increased (profit = $8,300 at capacity of 98). The optimization criterion, therefore, is seen to be an important factor in the analysis. It is conjectured that optimum capacity based on maximizing profit would be equal to the minimum peak daily demand because the facility would have a maximum utilization at that point. Capacities greater than this would result in unused spaces at least some of the day. However, this neglects ill will caused by inadequate size and, practically, may not be optimal.

DETERMINATION OF OPTIMUM LIFE-CYCLE CAPACITY

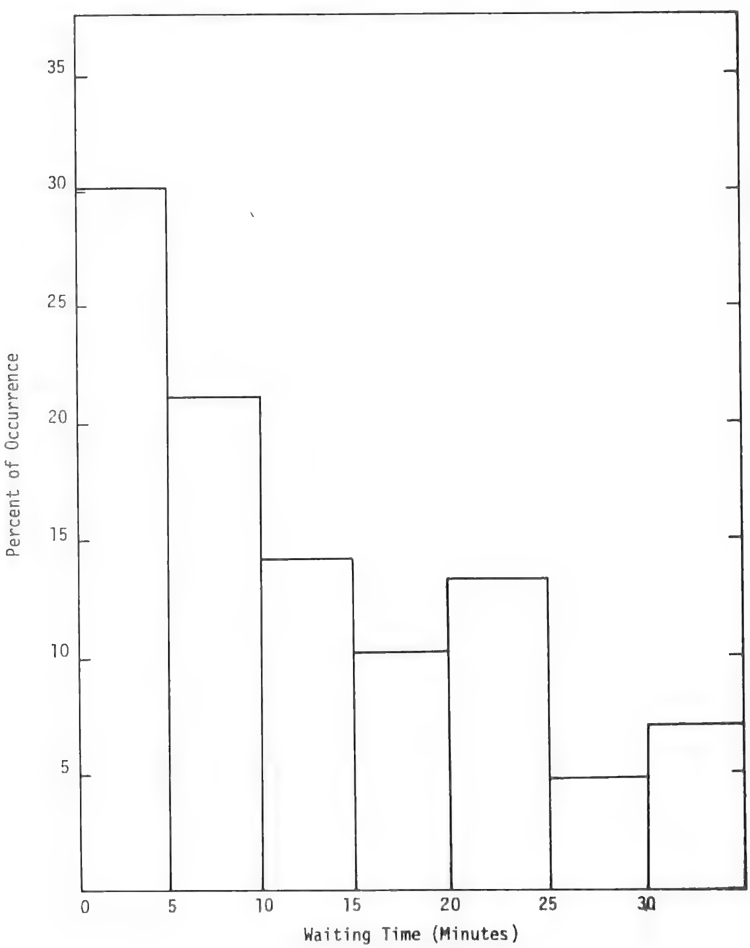
The preceding section determined the optimum capacity of a facility by using the demand characteristics of a 1-year period. Realistically, a terminal must be of optimum capacity over its entire life cycle, so the demand must be accurately projected

Table 1. Data for single-year analysis.

Capacity	Waiting Cost ($)	Terminal Cost[a] ($)	Total Cost[b] ($)	Wait Time (hour)	No. of Days Waiting Occurred	Maximum Queue Length	No. of Waiting Vehicles
60	87,298	9,500	96,798	32,016	245	120	27,053
70	32,750	10,250	43,000	12,015	199	87	16,080
80	8,852	11,000	19,852	3,264	125	61	7,096
90	1,627	11,750	13,378	602	54	40	2,012
94	736	12,050	12,786	272	32	35	1,065
95	591	12,125	12,716	219	28	33	898
96	468	12,200	12,668	174	26	31	775
97	375	12,275	12,650	139	23	29	635
98[c]	298	12,350	12,648	111	19	27	533
99	238	12,425	12,663	88	16	24	455
100	188	12,500	12,688	70	13	22	385
101	145	12,575	12,720	54	11	20	344
110	2.6	13,250	13,252	0.9	3	6	22
120	0	14,000	14,000	0	0	0	0
130	0	14,750	14,750	0	0	0	0

[a]Including $5,000 fixed cost. [b]Decision variable. [c]Optimum capacity.

Figure 6. Waiting time distribution.



over the life cycle to determine optimal life-cycle capacity. A yearly adjustment factor was applied to the demand rates, and the simulation period was extended over the life cycle of the system to determine optimum life-cycle capacity.

Although a unique optimum capacity may exist for each year's demand, yearly changes in terminal capacity may not be practical so that a fixed capacity must, in many cases, be used over the system life cycle. The optimal life-cycle capacity is the capacity that minimizes the present worth of the total cost over all years.

Several additional factors could be considered in this life-cycle analysis. The demand mix may change over future years as well as the variance of parking duration. The local economy (inflation, recession, and the like) must be evaluated over the future years. Inflation, for instance, would increase the value of waiting time as well as terminal operating costs. Accurate predictions are mandatory to finding the optimal solution.

The computer program developed during this study was used for this life-cycle analysis. It was assumed that the demand rate would increase by 6 percent per year (compound), while terminal fixed costs would remain constant at $5,000 per year and all other costs (waiting, terminal variable, and parking costs) would increase by 5 percent per year (compound). The trip-purpose mix within the total and the parking duration were assumed to remain the same as used previously. These are relatively simple assumptions concerning future changes but indicate quite drastic differences in the results. The terminal life cycle was assumed to be 10 years. Each capacity examined was simulated for each of the 10 years under the appropriate demand and cost structure. The present worth of the costs for each of the 10 years for each capacity was found, and the objective became that of finding the capacity that minimized the present worth total cost over the life cycle. Figure 7 shows this present worth function for various capacities and indicates that the optimum capacity under these assumptions was 146 spaces. Total cost increases rapidly for capacities less than optimum because of increased waiting time. Capacities greater than optimum reduce waiting and increase the level of service but at the expense of increased investment and terminal operating cost.

Data given in Table 2 show some interesting results for this optimum capacity. No waiting occurred for the first 5 years, but the waiting incurred in years 9 and 10 may be prohibitive. Profit was greater than terminal cost only in the last 4 years. This indicates that a better scheme than having a fixed capacity over the life cycle may be found. However, in many cases, a variable capacity may not be feasible, particularly with respect to land availability. Dynamic programming could be used to determine the optimum variable capacity program although it would prove to be an expensive method of analysis. Alternately, a constrained objective function could be employed to determine the optimum, minimum total cost capacity that would allow no more than a predetermined maximum number in a queue or a maximum amount of waiting time or both.

CONCLUSIONS AND RECOMMENDATIONS

The criteria on which the worth of the system is evaluated are principally economic in nature. That is, the "best" choice of a parking facility capacity for a given demand is simply the configuration that achieves the lowest cost solution in light of the economic trade-offs that are characteristic of such problems. The theoretical considerations should be basically incorporated into any practical application. If we are to illustrate the optimization concept, we should generalize the solution method by determining the capacity of an actual facility so that the method can be examined from both theoretical and practical levels.

The discrete-event simulation model developed by this study is a good method to determine the optimum capacity of minicar terminals and to analyze the effects of waiting with respect to users and the terminal area. The model provides a great deal of detailed information about the terminal system. Changes or additions to the program can be made easily to provide additional information that may be required for a particular application. The objective of other studies may be to maximize profits and determine return on investment rather than minimize costs, but the computer processing

Figure 7. Life cycle capacity costs.

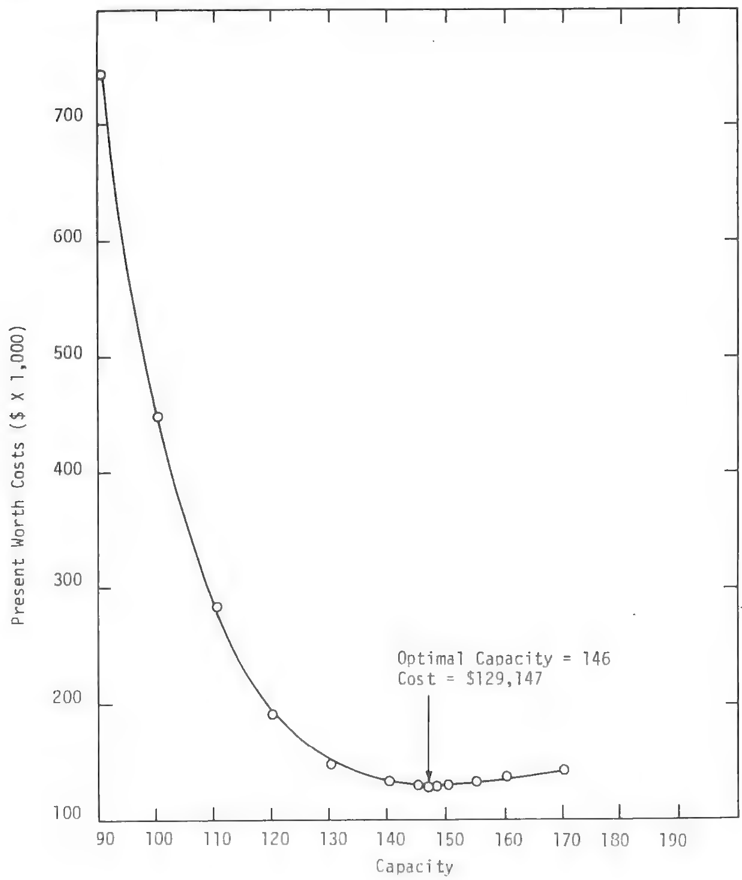


Table 2. Life cycle analysis—optimal capacity.

Year	Waiting Cost ($)	Terminal: Fixed Cost ($)	Terminal: Variable Cost ($)	Terminal: Total Cost ($)	Total Costs	No. of Days Waiting Occurred	No. of Waiting Vehicles	Maximum in Queue	Total Demand	Profit (disc.)	Waiting Time (min): Average	Waiting Time (min): Maximum
1	0	4,630	10,139	14,769	14,769	0	0	0	59,004	4,350	0	0
2	0	4,287	9,857	14,144	14,144	0	0	0	62,121	5,680	0	0
3	0	3,969	9,583	13,553	13,553	0	0	0	66,415	7,270	0	0
4	0	3,675	9,317	12,992	12,992	0	0	0	70,676	8,730	0	0
5	0	3,403	9,058	12,461	12,461	0	0	0	75,542	10,300	0	0
6	1.20	3,151	8,807	11,958	11,959	1	12	4	80,139	11,620	2.8	4.6
7	66.3	2,917	8,562	11,480	11,546	8	241	17	84,545	12,840	7.9	22
8	275	2,701	8,324	11,026	11,301	15	808	26	89,919	14,200	10	37
9	2,309	2,501	8,093	10,594	12,904	38	2,839	83	95,566	15,560	24.3	82
10	2,316	2,316	7,868	10,184	13,548	73	5,036	97	100,510	16,550	20.6	104
Total					129,147					107,100		

would remain basically the same. The most important factor in applying the model is to predict accurately the demand structure to be serviced. Although discussion of prediction models is beyond the scope of this paper, suffice it to say that the true stochastic nature of demand must be adequately predicted to obtain meaningful results.

It was also shown that optimization to maximize profits gave a very different solution than optimizing to minimize total costs. It is evident that the minimum-cost solution is superior because it provides for a larger terminal and thus a higher level of service to customers. In an actual application, it may be desirable to find the optimum capacity without studying the cost function over a wide range. The Fibonnaci search procedure (10) may be incorporated into the computer program to do this economically. This search technique guarantees that the optimal solution would be found in a minimum amount of computer processing time.

In an actual minicar system, an imbalance may exist between the daily arrival and departure rates at a given terminal. This would necessitate the redistribution of minicars sometime during the day or night to ensure the best distribution of vehicles throughout the system. An extension of this study, therefore, should include the possibility of having to wait for an available minicar at departure time and the redistribution interaction among terminals in the minicar network.

ACKNOWLEDGMENTS

This research was performed, in part, under a project supported by the National Science Foundation. The research contract was awarded to the Civil Engineering Department of Virginia Polytechnic Institute and State University.

REFERENCES

1. Altman, M., et al. Minicar Transit System: Feasibility Study, Books 1 and 2. Univ. of Pennsylvania, Philadelphia, 1968.
2. Baerwald, J. E., ed. Traffic Engineering Handbook. Institute of Traffic Engineers, 1965, pp. 459-500.
3. Capobianco, M. F. Using Probabilities to Compare Parking Disciplines. Traffic Quarterly, Vol. 22, No. 1, Jan. 1968, pp. 137-144.
4. Kole, C. P. Real-Time Parking Control. Traffic Engineering, April 1967, pp. 36-38.
5. Parzen, E. Stochastic Processes. Holden-Day, San Francisco, 1967, p. 118.
6. Pritskan, A. A. B., and Kiviat, P. J. Simulation With GASP. Prentice-Hall, Englewood Cliffs, N. J., 1968.
7. Schulman, L. L., and Stout, R. W. A Parking Study Through the Use of Origin-Destination Data. Highway Research Record 317, 1970, pp. 14-29.
8. Vuchic, V. R., ed. Minicar Transit System, Final Report. Univ. of Pennsylvania, Philadelphia, 1970.
9. Parking in the City Center. Wilbur Smith and Assoc., New Haven, Conn., 1965.
10. Wilde, D. J., and Beightler, C. S. Foundation of Optimization. Prentice-Hall, Englewood Cliffs, N. J., 1967.
11. Wilhelm, W. E. Determination of the Number, Location, and Capacity of Terminals in a Network. Virginia Polytechnic Institute, Master's thesis, 1970.
12. Wohl, M. A Methodology for Forecasting Peak and Off-Peak Travel Volumes. Highway Research Record 322, 1970, pp. 183-219.
13. Yu, J. C. A Parametric Analysis of Fleet Parking Terminal Capacity. Highway Research Record 317, 1970, pp. 30-39.
14. Yu, J. C., and Wilhelm, W. E. Optimization of Urban Transportation Terminal Network. Highway Research Record 348, 1971, pp. 94-108.
15. Yu, J. C., and Wilhelm, W. E. Optimal Location Patterns of Urban Transportation Terminals. Presented at the Joint Transportation Engineering Conference, Chicago, Oct. 1970.

SPONSORSHIP OF THIS RECORD

GROUP 1—TRANSPORTATION SYSTEMS PLANNING AND ADMINISTRATION
Charles V. Wootan, Texas A&M University, chairman

Committee on Transportation Information Systems and Data Requirements
Robert E. Barraclough, Information Systems Consultant, chairman
Ben Alexander, Geoffrey J. H. Brown, Henry W. Bruck, Edward N. Dodson, Kenneth J. Dueker, David J. Dunlap, Gideon Hashimshony, George T. Lathrop, Joseph M. Manning, Alfred J. Marston, Peter S. Parsonson, Albin Pearson, Alan E. Pisarski, Sydney R. Robertson, George N. Sarames, Arthur B. Sosslau, D. L. Wieman, J. E. Wright

Task Force on Application of Economic Analysis to Transportation Problems
Robley Winfrey, consulting civil engineer, chairman
James B. Chiles, Paul J. Claffey, Donald O. Covault, Dan G. Haney, Herbert Mohring, Roger G. Noll, Clarkson H. Oglesby, Jan De Weille

James A. Scott and Kenneth E. Cook, Highway Research Board staff

The sponsor is identified by a footnote on the first page of each report.

THE National Academy of Sciences is a private, honorary organization of more than 800 scientists and engineers elected on the basis of outstanding contributions to knowledge. Established by a congressional act of incorporation signed by Abraham Lincoln on March 3, 1863, and supported by private and public funds, the Academy works to further science and its use for the general welfare by bringing together the most qualified individuals to deal with scientific and technological problems of broad significance.

Under the terms of its congressional charter, the Academy is also called upon to act as an official—yet independent—adviser to the federal government in any matter of science and technology. This provision accounts for the close ties that have always existed between the Academy and the government, although the Academy is not a governmental agency and its activities are not limited to those on behalf of the government.

The **National Academy of Engineering** was established on December 5, 1964. On that date the Council of the National Academy of Sciences, under the authority of its act of incorporation, adopted articles of organization bringing the National Academy of Engineering into being, independent and autonomous in its organization and the election of its members, and closely coordinated with the National Academy of Sciences in its advisory activities. The two Academies join in the furtherance of science and engineering and share the responsibility of advising the federal government, upon request, on any subject of science or technology.

The **National Research Council** was organized as an agency of the National Academy of Sciences in 1916, at the request of President Wilson, to provide a broader participation by American scientists and engineers in the work of the Academy in service to science and the nation. Its members, who receive their appointments from the President of the National Academy of Sciences, are drawn from academic, industrial, and government organizations throughout the country. The National Research Council serves both Academies in the discharge of their responsibilities. Supported by private and public contributions, grants, and contracts and by voluntary contributions of time and effort by several thousand of the nation's leading scientists and engineers, the Academies and their Research Council thus work to serve the national interest, to foster the sound development of science and engineering, and to promote their effective application for the benefit of society.

The **Division of Engineering** is one of the eight major divisions into which the National Research Council is organized for the conduct of its work. Its membership includes representatives of the nation's leading technical societies as well as a number of members-at-large. Its Chairman is appointed by the Council of the Academy of Sciences upon nomination by the Council of the Academy of Engineering.

The **Highway Research Board** is an agency of the Division of Engineering. The Board was established November 11, 1920, under the auspices of the National Research Council as a cooperative organization of the highway technologists of America. The purpose of the Board is to advance knowledge of the nature and performance of transportation systems through the stimulation of research and dissemination of information derived therefrom. It is supported in this effort by the state highway departments, the U.S. Department of Transportation, and many other organizations interested in the development of transportation.

HIGHWAY RESEARCH BOARD
NATIONAL ACADEMY OF SCIENCES—NATIONAL RESEARCH COUNCIL
2101 Constitution Avenue Washington, D. C. 20418

ADDRESS CORRECTION REQUESTED

NON-PROFIT ORG.
U.S. POSTAGE
PAID
WASHINGTON, D.C.
PERMIT NO. 42970

001033 [redacted]
[redacted] GINEERING
[redacted] G
[redacted] 78712